



enacomp

2013

X Encontro Anual de Computação

Tema:
Computação,
da teoria à prática

Anais

ISSN 2178-6992

REALIZAÇÃO:



PROEC
PRÓ-REITORIA DE EXTENSÃO E CULTURA



CECAC
COORDENAÇÃO DE EXTENSÃO
E CULTURA DO CAMPUS CATALÃO

SUORTE:



FAPEG
FUNDAÇÃO DE AMPARO
À PESQUISA
DO ESTADO DE GOIÁS

APOIO:



Ficha catalográfica

EnAComp, X Encontro Anual de Computação

E56a Anais do X Encontro Anual de Computação - EnAComp
2013 / Luanna Lopes Lobato e Thiago Jabur Bittar -
Organizadores. -- Catalão, 2013.
410 p.

Anais contendo trabalhos completos e resumidos,
2013, março.

1. Computação. 2. Informática. 3. Ciência da
Computação. I. Lobato, Luanna Lopes e Bittar, Thiago
Jabur. organizadores. Autor coletivo (autores de cada
trabalho indicados).

Comissão Organizadora

Coordenação

Profa. Dra. Luanna Lopes Lobato

Vice-Coordenação

Prof. Dr. Thiago Jabur Bittar

Secretário

Humberto Lidio Antonelli

Apoio

Prof. Acrísio José do Nascimento Júnior

Alunos

Anelisa Pereira da Silva
César Augusto Verçosa Mariano
Cleriston Aparecido Dantas
Danilo Martins Silvério Cardoso
Diego Martins de Almeida
Elizabeth Cristina Nascimento Fortunato
Fernanda Bontempo Faria
Humberto Lidio Antonelli
Iohan Gonçalves Vargas
Lorena Aparecida Rezende
Lucas Ângelo da Silveira
Luiz Gustavo Dias
Nilton Mendes de Souza
Paulo Henrique Ferreira Fontoura Junior
Thiago de Almeida Bastos
Victor Portilho

Comissão Científica

Alexsandro Santos Soares (UFU/FC – Uberlândia/MG)
Carla Ilane Moreira Bezerra (UFC – Quixadá/CE)
Celso Augusto Raposo Lisboa Brennand (UFRPE/DCC – Serra Talhada/PE)
Ciro Meneses Santos (UFVJM – Teófilo Otoni/MG)
Crescencio Rodrigues Lima Neto (IFBA/Info – Vitória da Conquista/BA)
Cristiane Akemi Yagunuma (UFSCar/DC – São Carlos/SP)
Cristiane de Fátima dos Santos Cardoso (IFGO – Urutaí/GO)
Dalton Matsuo Tavares (UFG/DCC – Catalão/GO)
Daniele Santini Jacinto (USP/IFSC – São Carlos/SP)
Elias Adriano Nogueira da Silva (USP/ICMC – São Carlos/SP)
Ednaldo Dilonzo de Souza Filho (Dataprev – Recife/PE)
Felipe Domingos da Cunha (UFMG/DCC – Belo Horizonte/MG)
Fernando Antônio Asevedo Nóbrega (USP/ICMC – São Carlos/SP)
Gleibson Rodrigo Silva de Oliveira (UFPE/CIn – Recife/PE)
Gláucia Carreiro Boechat (UNICAMP – Campinas/SP)
Henrique Emanuel Mostaert Rebêlo (UFPE/CIn – Recife/PE)
Hugo Sica de Andrade (UFBA/DCC – Salvador/BA)
Iuri Santos Souza (RiSE – Salvador/BA)
Ivan do Carmo Machado (UFBA/DCC – Salvador/BA)
Ivan Sedin (UFG/DCC - Catalão/GO)
Ivonei Freitas da Silva (UNIOESTE/DCC – Cascavel/PR)
Jeneffer Cristine Ferreira (UFRPE – Recife/PE)
João Guilherme Maia de Menezes (UFMG/DCC – Belo Horizonte/MG)
Joel da Silva (UFSM/CAFW – Santa Maria/RS)
Jonatas Bastos (IFBA/Info – Irecê/BA)
José Jorge Lima Dias Junior (UFPB/DCE – Rio Tinto/PB)
Juliana de Albuquerque Gonçalves Saraiva (UFRPE/UAG – Garanhuns/PE)
Larissa Rocha Soares (UFBA/DCC – Salvador/BA)
Leandro Agostini do Amaral (USP/ICMC – São Carlos/SP)
Luanna Lopes Lobato (UFG/DCC – Catalão/GO)
Marcelo Gabriel Almiron (UFMG/DCC – Belo Horizonte/MG | UNIGE, CUI, Genebra)
Márcio Antônio Duarte (UFG/DCC – Catalão/GO)
Márcio de Souza Dias (UFG/DCC – Catalão/GO)
Marcos Antônio de Matos Laia (USP/ICMC – São Carlos/SP)
Maria Isabel Oliveira (UFBA/IF – Salvador/BA)
Mehran Misaghi (Sociesc - Joinville/SC)
Muriel de Souza Godoi (UTFPR – Apucarana/PR)
Nádia Félix Felipe da Silva (USP/ICMC – São Carlos/SP)
Rafael de Amorim Silva (UFAL/IC – Maceió/AL)

Raphael Pereira de Oliveira (UFBA/DCC – Salvador/BA)
Rhavy Maia Guedes (IFPB/DE – Campina Grande/PB)
Renata Maria de Souza (CESAR – Recife/PE)
Renato Lima Novais (IFBA – Santo Amaro/BA)
Robson Eduardo de Grande (University of Ottawa – Canadá)
Roberto Mendes Finzi Neto (UFU/FEMEC - Uberlândia)
Roberto Rigolin Ferreira Lopes (USP/ICMC – São Carlos/SP)
Rodrigo Rocha Gomes e Souza (UFBA/DCC – Salvador/BA)
Selma Terezinha Milagre (FEE/UFU – Uberlândia/MG)
Sérgio Francisco da Silva (UFG – Catalão/GO)
Sérgio Francisco Tavares de Oliveira Mendonça (UFRPE – Garanhuns/PE)
Simone da Silva Amorim (UFBA/DCC – Salvador/BA)
Stella Jacyszyn Bachega (UFG – Catalão/GO)
Tássio Ferreira Vale (RiSE – Salvador/BA)
Thiago Souto Mendes (IFBA – Campus Santo Amaro/BA)
Thiago Jabur Bittar (UFG/DCC – Catalão/GO)
Thiago Henrique Burgos de Oliveira (Fraunhofer/IESE – Alemanha)
Ulysses Roberto Chaves Vitor (UFPE/DEE – Recife/PE)
Valdirene Fontanette (Itera – São Carlos/SP)
Yguaratã C. Cavalcanti (SERPRO – Florianópolis/SC)
Wesley Barbosa Thereza (UNEMAT/DC – Mato Grosso)
Wylliams Barbosa Santos (UFRPE – Recife/PE)

Entidades Envolvidas

Realização

Departamento de Ciência da Computação (DCC)

Universidade Federal de Goiás (UFG)

Apoio

Sociedade Brasileira de Computação (SBC)

Grupo ReseArch

Suporte

CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

FAPEG - Fundação de Amparo à Pesquisa do Estado de Goiás

Prefácio

Este livro de Anais contém os artigos aceitos para apresentação no X Encontro Anual de Computação – EnAComp 2013, realizado na Universidade Federal de Goiás (UFG), campus Catalão, durante os dias 12 a 15 de março de 2013.

O EnAComp é um evento que tem por objetivo reunir acadêmicos e profissionais envolvidos com a Computação e áreas afins, proporcionando a troca de conhecimento, experiências e apresentando as novas perspectivas referentes à Computação, assim como facilitando a discussão de inovações tecnológicas.

O tema deste ano é “*Computação, da teoria à prática*”. A ideia que envolveu a escolha deste tema está calcada na necessidade de aproximarmos sempre as teorias ensinadas na universidade à prática implementada nas indústrias, uma vez que os alunos devem ser preparados para atuar na academia e no mercado de trabalho.

A Comissão Científica do EnAComp 2013 é formada por um grupo multidisciplinar de especialistas relacionados à Computação, constituído por mais de 60 pesquisadores, de diferentes instituições e empresas do país, aos quais foi atribuído a responsabilidade de avaliar, em um processo de revisão as cegas, os trabalhos para cada uma das áreas de interesse relacionadas ao evento.

No EnAComp 2013 foram recebidos cerca de 130 trabalhos, em forma de artigos completos e pôsteres, sobre pesquisas que tem sido desenvolvidas na academia e na indústria. Destes, 57 artigos foram aceitos para publicação, sendo 35 orais e 22 pôsteres. A publicação dos anais é disponibilizada em formato eletrônico (na Internet e em CD), com o número de ISSN 2178-6992.

Agradecemos a todos que, direta ou indiretamente, tem colaborado positivamente para a realização do EnAComp 2013 (alunos envolvidos na comissão organizadora, palestrantes, autores, comitê científico, patrocinadores, instituições e familiares). Fizemos nosso melhor para oferecer um interessante encontro, bem como para proporcionar aos participantes ótimos dias durante a estadia em Catalão, GO.

Catalão, 12 de março de 2013.

Coordenação do X EnAComp – 2013

Profa. Dra. Luanna Lopes Lobato

Prof. Dr. Thiago Jabur Bittar

Sumário

<i>Reflexões sobre Síntese Automática de Sistemas de Informação Temporais através de Modelos..</i>	<i>1</i>
<i>Avaliação de Desempenho de Reconhecimento Automático de Faces Utilizando as Técnicas de Análise de Componentes Principais e de Fisherfaces</i>	<i>9</i>
<i>Análise de Segurança da Camada de Notificação do SparkleShare</i>	<i>17</i>
<i>Estimação do Tempo Médio de Resposta do Servidor em Plataformas de Ensino a Distância do Tipo Cliente-Servidor Utilizando Modelo Baseado em Teoria de Filas</i>	<i>25</i>
<i>Acesso Remoto a Redes de Sensores Sem Fio Usando VPN.....</i>	<i>33</i>
<i>O Uso da Tecnologia de Informação em Sala de Aula: Um Estudo de Caso.....</i>	<i>41</i>
<i>Monitoramento de Paisagens Urbanas com Redes de Sensores para Auxílio a Decisão em Gestão Participativa.....</i>	<i>49</i>
<i>Simulação baseada em autômatos celulares para a evacuação emergencial de pedestres em ambientes com geometria complexa.....</i>	<i>57</i>
<i>Uma abordagem de código único para aplicações independentes de provedor de bases de dados relacionais.....</i>	<i>65</i>
<i>Metodologia de análise para migração de sistemas de médio e grande portes para a nuvem no modelo SaaS</i>	<i>73</i>
<i>Utilização de dispositivos móveis para o monitoramento e controle de transdutores inteligentes aplicados na automação residencial.....</i>	<i>81</i>
<i>Desempenho do Uso de Visões nas Consultas em Diferentes Servidores de Banco de Dados Relacionais</i>	<i>89</i>
<i>Ambiente Virtual de Cenário Urbano para Simulação Multiagentes - Experiência do Projeto ADGEPA.....</i>	<i>97</i>
<i>FlexTest – Um Framework Flexível para a Automação de Testes</i>	<i>105</i>
<i>Convolução como Caso de Estudo para Formalização em PVS da Correção Funcional de Implementações em FPGA's.....</i>	<i>113</i>
<i>Mecanismo de Controle de Fluxo Aplicado na Transmissão de Vídeo no Sistema LTE.....</i>	<i>121</i>
<i>Aplicação de Processamento Paralelo em método iterativo para solução de sistemas lineares</i>	<i>129</i>
<i>Em Direção a uma Ferramenta de Suporte a Revisão Sistemática de Literatura.....</i>	<i>137</i>
<i>Construção de uma Base de Dados para Auxiliar a Avaliação de Sistemas de Detecção de Intrusos em uma Rede IEEE 802.11 com Criptografia WEP, WPA e WPA2 Habilitada.....</i>	<i>145</i>
<i>Aplicativo classificador de imagens de notas de real para plataforma Android</i>	<i>152</i>
<i>Aplicação para dispositivos móveis utilizando tecnologias interativas: A Realidade Virtual e Aumentada aplicada ao estudo da anatomia humana.....</i>	<i>160</i>
<i>Aplicação de ferramentas automatizadas para geração de formulários Web acessíveis.....</i>	<i>168</i>

<i>Otimização das Bandas Fotônicas Proibidas Absolutas de Cristais Fotônicos Anisotrópicos Bidimensionais utilizando Sistemas Imunológicos Artificiais</i>	<i>176</i>
<i>SEDOR - Sistema Especialista para Diagnóstico Odontológico baseado em Radiografias.....</i>	<i>184</i>
<i>Avaliação do Desempenho de Esquemas de Cruzamento em Algoritmos Genéticos na Otimização de Bandas Proibidas em Cristais Fotônicos</i>	<i>192</i>
<i>Avaliação da Usabilidade em Sistemas de Gerenciamento de Biblioteca: Software SIGA e SophiA.....</i>	<i>200</i>
<i>Padrões e diretrizes arquiteturais para escalabilidade.....</i>	<i>208</i>
<i>Otimização de circuitos usando Simulated Annealing</i>	<i>216</i>
<i>Considerações sobre um desenvolvimento acessível por meio do registro de Design Rationale</i>	<i>222</i>
<i>Autenticação Biométrica em Redes de Telecomunicação Aeronáutica.....</i>	<i>230</i>
<i>Reconfiguração de Redes de Distribuição de Energia Elétrica Utilizando Metaheurística Colônia de Formigas Modificado.....</i>	<i>238</i>
<i>Uma proposta de desenvolvimento de um sistema de software para apoio ao gerenciamento de doenças relacionadas à Leishmaniose</i>	<i>246</i>
<i>Um breve estudo sobre Gerenciamento de Riscos em Projetos de Software no Brasil.....</i>	<i>252</i>
<i>Concepção de uma Ferramenta Parser para Extração de dados do metamodelo ECore.....</i>	<i>260</i>
<i>Aplicando Padrões de Seleção no Ensino de Programação de Computadores para Estudantes do Primeiro Ano do Ensino Médio Integrado</i>	<i>268</i>
<i>Modelagem de Vídeo MPEG-4 Utilizando Cascata Multifractal com Multiplicadores Autorregressivos.....</i>	<i>276</i>
<i>Caracterização de QoS para Conteúdo Multimídia em Rede VPN utilizando Rede Neural MLP</i>	<i>284</i>
<i>A TI como uma Ferramenta de Gestão Estratégica nas Operações de Uma Empresa do Setor de Logística.....</i>	<i>290</i>
<i>Os desafios para a mobilização de aplicações baseadas em plataforma Web.....</i>	<i>294</i>
<i>Lógica para Semáforo Inteligente Baseado na Mineração de Dados por Algoritmo Genético Transgênico</i>	<i>301</i>
<i>Motor de inferências em plataforma OSGi.....</i>	<i>309</i>
<i>Um Sistema para Visualização Tridimensional das Mamas a partir da Curva do Limite Inferior Frontal das Mamas</i>	<i>313</i>
<i>Deteção, caracterização e recuperação de faces</i>	<i>317</i>
<i>Estudo de Caso Usando o Framework Robot Operating System (ROS)</i>	<i>324</i>
<i>Desenvolvimento de um Simulador para o Treinamento de Operações para Profissionais do Setor Elétrico.....</i>	<i>332</i>
<i>Construção de Software capaz de traçar o Perfil Nutricional e Antropométrico dos estudantes no IFPB campus Campina Grande.....</i>	<i>337</i>
<i>Estudo de Caso Usando a Plataforma Microsoft® Robotics Developer Studio (RDS).....</i>	<i>341</i>

<i>Estudo de Caso Usando o Framework LeJOS.....</i>	<i>349</i>
<i>Preditor de números para lotofácil: uma abordagem usando algoritmos evolutivos</i>	<i>357</i>
<i>Interdisciplinaridades entre Pesquisa Operacional e Ciência da Computação.....</i>	<i>363</i>
<i>OpenNI e suas Aplicações</i>	<i>371</i>
<i>SELECTRESS – Sistema Web aplicado ao melhoramento genético vegetal para resistência a fatores ambientais estressantes</i>	<i>379</i>
<i>Modelagem Multifractal BetaMWM Adaptativa para Tráfego de Redes de Computadores.....</i>	<i>383</i>
<i>Um Sistema Móvel de Compras Rápido e Seguro via QR Code e Vitrine Virtual.....</i>	<i>391</i>
<i>Resolução de PPLs através dos métodos Simplex e Simplex em Duas Fases</i>	<i>395</i>
<i>Gerenciamento de melhorias de informações utilizando Data Warehouse.....</i>	<i>399</i>
<i>Application programming interface Library para Algoritmo Genético com Operador Transgênico</i>	<i>407</i>

Reflexões sobre Síntese Automática de Sistemas de Informação Temporais através de Modelos

Valdemar Vicente Graciano Neto

¹Instituto Federal de Educação, Ciência e Tecnologia de Goiás (IFG) - Câmpus Formosa
Rua 64, esq. c/ Rua 11, s/n, Expansão Parque Lago – 73.813-816. Formosa - GO – Brasil

vgracianoneto@gmail.com

Resumo. *Software de Sistemas de Informação (SI) naturalmente precisam armazenar informações temporais. Para tanto, é necessário que o banco de dados seja do tipo temporal. Entretanto, os sistemas gerenciadores de banco de dados modernos dão suporte a apenas poucos tipos de aspectos temporais. Em virtude disso, é comumente necessário um componente de software que administre a temporalidade do software de SI e a temporalidade do próprio banco de dados. Este artigo apresenta reflexões e decisões de projeto que servirão de insumo para conceber uma arquitetura dirigida por modelos para síntese automática de SI com suporte a temporalidade.*

1. Introdução

O desenvolvimento de software de Sistemas de Informação (SI) mantém-se ainda como uma atividade cara e complexa. Grandes esforços são demandados em todos os estágios do processo de desenvolvimento para lidar com a qualidade dos requisitos do cliente dentro de orçamentos e prazos reduzidos.

Uma complexidade notável com a qual é preciso lidar é o tratamento de aspectos temporais em SI. Ela surge naturalmente em várias categorias de software de SI e consiste na necessidade de manter informações que são alteradas ao longo do tempo.

SI são naturalmente associados a Bancos de Dados (BD) e os Sistemas Gerenciadores de Banco de Dados (SGBD) modernos oferecem poucos recursos para tratamento de temporalidade. Georg concluiu, a partir de um estudo conduzido sobre temporalidade nos SGBD Objeto-Relacionais Oracle e PostgreSQL, que, apesar de não impedir a aplicação de temporalidade em BD, essas tecnologias não dão o suporte necessário para fazê-lo de forma transparente e eficiente [de Castro Georg 2010].

O Desenvolvimento de Software Dirigido por Modelos (DSDM) [Stahl et al. 2006] é uma forma de lidar com as complexidades do processo de desenvolvimento de SI ao substituir o código-fonte por modelos abstratos que representam especificações em alto-nível de abstração.

O código é gerado automaticamente a partir dos modelos de modo que o código não demande manipulação direta, atividade onerosa e propensa a erros.

Este artigo apresenta resultados preliminares, reflexões e decisões de projeto que servirão de insumo para conceber uma arquitetura dirigida por modelos para geração automática de aspectos de gerência de temporalidade em projetos de desenvolvimento de software de SI. Para tratar deste tema, a Seção 2 discute o embasamento teórico desta

pesquisa; a Seção 3 discute decisões de projeto para arquitetura de uma ferramenta de transformação de modelos para síntese de SI temporal; e a Seção 4 apresenta as conclusões e trabalhos futuros.

2. Embasamento Teórico

Esta seção apresenta os pilares teóricos sobre os quais esta pesquisa se apoia.

2.1. Temporalidade em SI e BD

Sistemas de Informação (SI) de grande porte são inerentemente temporais. Isto é, necessitam tratar a evolução sofrida pelos dados e guardar o histórico de modificações ocorridas.

Como as informações de um SI são comumente gravadas em um Banco de Dados, é necessário prover modos de tratar a temporalidade em nível de aplicação, uma vez que os Sistemas Gerenciadores de Banco de Dados (SGBD) modernos não dão vasto suporte ao tratamento de temporalidade [de Castro Georg 2010].

Bancos de Dados (BD) convencionais representam objetos do mundo real em um único momento do tempo, o atual. BD temporais, por sua vez, permitem que se pesquise por estados anteriores (passado) ou posteriores (futuro), utilizando uma linguagem de consulta temporal. Por estado futuro deve-se entender que a informação já está presente no banco de dados mas somente será válida a partir de um instante futuro [de Castro Georg 2010], como quando um débito em uma conta ocorre num final de semana, mas só será efetivado no primeiro dia útil seguinte.

Para tanto, porções de código são acrescentadas à arquitetura do SI bem como ao código de manipulações de dados do BD (*stored procedures*) para gerenciar a temporalidade do SI.

Um modelo de dados convencional de BD apresenta apenas duas dimensões: registro (tupla) e campo. Neste caso, a mudança de qualquer valor de instância implica na perda do valor anterior do campo no qual o dado foi modificado.

Os modelos temporais acrescentam uma terceira dimensão: o tempo. Esta dimensão atribui uma informação temporal a cada valor. Desta forma, a mudança de qualquer valor não inutiliza o valor anterior.

Em aplicações de BD utiliza-se diversos tipos de dados temporais com diferentes graus de precisão. Para tanto, em BD temporais estabelece-se os conceitos de *tempo de validade* e *tempo de transação*.

Tempo válido (ou de validade) é o tempo em que o valor associado a ele é válido no mundo real. É um valor que é explicitamente informado pelo usuário. Ou seja, o usuário tem total controle sobre o tempo de validade de uma informação. Isso demanda o acréscimo de dois campos a cada tabela: TVI e TVF que representam, respectivamente, o tempo válido inicial e o tempo válido final.

Tempo de transação indica o instante em que os valores foram efetivados no BDT. A informação associada ao tempo de transação será válida (atual no banco de dados) enquanto não for excluída logicamente. Este tipo de informação temporal é inacessível ao usuário, pois é gerada e gerenciada pelo SGBD. Consequentemente, o tempo de transação

não pode ser posterior ao tempo atual do sistema. Um exemplo de controle de tempo de transação é o uso do tipo de dados **TIMESTAMP** em logs de operações do SGBD.

Tanto em BD de tempo de transação quanto em BD de tempo válido, uma alteração gera exatamente uma nova tupla.

Logo, em virtude de uma quantidade elevada de tuplas em uma mesma tabela e devido ao fato de que nem todas elas representam dados válidos no momento presente, é necessário filtrar os dados todas as vezes que else são disponibilizados para o usuário. Assim, é necessário gerar funcionalidades de software que gerenciem tanto a alteração dos dados no banco de dados de modo a possibilitar a atualização dos dados temporais quanto a leitura dos dados em virtude de consultas realizadas a estes tipos especiais de BD.

Em virtude da necessidade de gerar código manualmente e acrescentá-lo na arquitetura do SI (algo repetitivo e propenso a erros), vislumbrou-se a aplicação de princípios de Desenvolvimento de Software Dirigido por Modelos (DSDM) para a geração destas porções de software a partir de modelos em alto-nível de abstração, tema discutido na próxima seção.

2.2. Desenvolvimento de Software Dirigido por Modelos

Existem modelos específicos para cada fase do processo de desenvolvimento de software. O modelo de requisitos é usado como entrada para discussão e produção de modelos de projeto e arquitetura. Estes modelos são levados em consideração para a estruturação do código-fonte e o código-fonte é o insumo para especificação de casos de teste. O processo de desenvolvimento de software resume-se a uma série de *transformações entre modelos* conduzidas majoritariamente de modo manual [Graciano Neto et al. 2010].

DSDM quebra esse paradigma por automatizar este processo de transformação. O objetivo deste novo modelo prescritivo de processo de desenvolvimento de software é elevar o nível de abstração na produção de software trazendo ganhos como geração automática de código, produtividade, manutenibilidade, portabilidade, reusabilidade e rastreabilidade [Kleppe et al. 2003, Miller and Mukerji 2003, Graciano Neto et al. 2010].

Para tanto é necessário um software transformador de modelos e uma definição de transformação - um conjunto de regras que especificam como um modelo de origem pode ser transformado em um modelo destino (código-fonte) [Kleppe et al. 2003]. Estas regras ou definições de transformação devem ser plugáveis à ferramenta de transformação.

Para tornar a automação uma realidade, os modelos precisam ter um significado bem definido [Mellor et al. 2003]. Uma forma de prover tal significado é criar *metamodelos*, ou seja, modelos que fazem asserções sobre o que pode ser expresso em outros modelos [Seidewitz 2003]. Além disso, eles proveem sintaxe e semântica definidas para os modelos produzidos a partir deles.

A especificação MDA (*Model-Driven Architect*) - especificação seminal do OMG para guiar a produção de software segundo DSDM - não apresenta propostas arquiteturais para as ferramentas de transformação dirigidas por modelos. Logo, cada iniciativa implementa a transformação de uma forma diferente.

A MDA propõe que um *Modelo Independente de Plataforma* (*Platform Inde-*

pendent Model - PIM) - como um modelo de domínio - seja usado como entrada para uma ferramenta de transformação para produzir automaticamente um *Modelo Específico de Plataforma (Platform Specific Model - PSM)* - como um modelo arquitetural ou de projeto, ou código-fonte.

A próxima seção mostra como estas ideias se relacionam para estabelecer uma arquitetura de uma ferramenta capaz de sintetizar software de SI com suporte à temporalidade.

3. Decisões de Projeto para a Ferramenta de Síntese de SI

Os modelos utilizados para gerar software de SI representam, em geral, os três principais aspectos de um SI: domínio, interação e processo de negócio [da Costa et al. 2010, Loja et al. 2010, de Oliveira et al. 2011]. O domínio, entretanto, permite gerar um SI completo uma vez que, a partir dele, é possível gerar o BD, os objetos de negócio, as funcionalidades CRUD (Criar, Ler, Atualizar e Deletar) de persistência, e as telas para interação [Almeida et al. 2009].

Logo, pela suficiência do modelo de domínio como modelo de origem, por decisão de projeto, optou-se por utilizar apenas este modelo como modelo de entrada para a criação de um SI completo temporal. Da Costa et al. apresentaram um metamodelo para geração automática de Sistemas de Informação Empresariais [da Costa et al. 2010]. Este metamodelo foi ampliado e estendido para acrescentar características de domínio que são necessárias para a geração de um SI completo. Esta extensão é apresentada em [Graciano Neto 2012] mas não é reproduzida aqui por questões de espaço.

Para a síntese de um SI com características temporais através de um processo de DSMD, foram tomadas as seguintes decisões de projeto:

1. **O SI temporal (e consequentemente o BD) gerado será dotado de tempo de validade e tempo de transação;**
2. **O padrão arquitetural adotado para o SI temporal gerado é o de Camadas** [Fowler 2002];
3. **Todos os elementos do metamodelo de origem (metadados e respectivos atributos) serão dotados de tempo inicial e final de validade e transação.** Isso implica na adição de quatro campos às tabelas geradas a partir de cada um destes metadados;
4. **O metamodelo de domínio (origem) a partir do qual deriva-se o modelo de domínio de origem é especificado em XML.** [Graciano Neto 2012] mostra uma versão XML completa do metamodelo. Um trecho deste metamodelo instanciado é mostrado à frente;
5. **O paradigma de geração do SI temporal é o *preenchimento de template de código*** [Wittman 2010]. Isto é, a partir da definição XML do modelo de entrada, são extraídas as informações para preencher templates de código que integrarão o SI temporal final. Os *scripts* de geração do BD temporal, as funcionalidades de gerência de temporalidade, e o CRUD da aplicação são gerados através do preenchimento de *templates* de código. A geração da interface é feita sob o paradigma IF-ELSE [Ma and Yang 2010, Graciano Neto and de Oliveira 2011], isto é, para cada tipo de dado de entrada, um tipo específico de elemento gráfico de interação é gerado (rótulo, campo de texto, etc.);

6. **A temporalidade é inserida na ferramenta de transformação, não no meta-modelo de origem.** Temporalidade em SI não é um assunto de domínio de grande parte dos engenheiros de software. Além disso, seriam necessários trechos de código referentes a temporalidade tanto no metamodelo quanto na ferramenta de transformação de modelos. Para evitar problemas de manutenção e evolução da ferramenta de transformação de modelos, optou-se por manter as questões de geração de temporalidade apenas na ferramenta de transformação;
7. **A geração de código dirigida por modelos será adotada em grande parte do processo.** Na geração das funcionalidades CRUD da camada de aplicação, a geração de código será realizada de modo convencional por ser uma funcionalidade padrão para qualquer tipo de SI e não exige informações específicas do modelo de origem;
8. **O padrão arquitetural para trânsito de dados dentro da arquitetura do SI gerado é o padrão de *Transfer Object* [Fowler 2002];**
9. **A geração de BD temporais inicialmente será voltada ao SGBD Postgres;**
10. ***Stored procedures* serão geradas para realizar o CRUD temporal dentro do BD.** Isto é, as atividades básicas de um banco de dados como Criar, Ler, Atualizar e Deletar valores serão invocados pela camada de Persistência uma vez que são necessárias manipulações nas tuplas de cada tabela;

As decisões de projeto 2, 8 e 9 são herdadas de uma versão anterior de um framework dirigido por modelos para geração de SI [Almeida et al. 2009] sobre o qual esta ferramenta se baseia.

Para ilustrar o processo de geração do SI temporal, o Código 1 traz um trecho de entrada do metamodelo em XML com a especificação de uma Entidade Forte chamada Pessoa Física que possui dois atributos alfanuméricos: cpf e nome. Dentro da tag *entidade* consta o nome da entidade em questão, um valor booleano que representa se a entidade é forte ou fraca, e o mnemônico desta entidade, que trata-se de um identificador alfanumérico único que é utilizado para dar nome à tabela gerada no banco de dados.

Esta entidade é um metadado de negócio. Ela possui dois atributos. Cada atributo possui um mnemônico (que dará nome ao campo da tabela gerada), cardinalidades mínima e máxima (referentes à quantidade de valores que aquele campo pode receber), um valor booleano que diz se aquele campo será parte de chave, e outro que diz se os valores daquele campo devem ser únicos. O atributo CPF possui um tamanho de 11 caracteres e não possui valor *default*. O campo 'nome' não é único e nem parte de chave, possui um tamanho de 50 caracteres e não possui valor *default*.

Code 1. Trecho do modelo de entrada em XML.

```

1 <?xml version="1.0" encoding="iso-8859-1"?>
2 <aplicacaoSI>
3 <metadadoNegocio>
4   <entidade>
5     <nomeEntidade>Pessoa Física</nomeEntidade>
6     <entidadeForte>true</entidadeForte>
7     <mnemonico>PesFis</mnemonico>
8   </entidade>
9   <atributos>
10    <atributo>
```

```

11      <mnemonico>cpf</mnemonico>
12      <cardinalidadeMaxima>1</cardinalidadeMaxima>
13      <cardinalidadeMinima>1</cardinalidadeMinima>
14      <ehParteChave>true</ehParteChave>
15      <ehUnico>true</ehUnico> <!-- TRUE ou FALSE -->
16      <!-- ALFANUMERICO -->
17      <tamanhoAlfanumerico>11</tamanhoAlfanumerico>
18      <valorDefaultAlfanumerico></valorDefaultAlfanumerico>
19    </atributo>
20    <atributo>
21      <atributo>
22        <mnemonico>nome</mnemonico>
23        <cardinalidadeMaxima>1</cardinalidadeMaxima>
24        <cardinalidadeMinima>1</cardinalidadeMinima>
25        <ehParteChave>false</ehParteChave>
26        <ehUnico>false</ehUnico> <!-- TRUE ou FALSE -->
27        <!-- ALFANUMERICO -->
28        <tamanhoAlfanumerico>50</tamanhoAlfanumerico>
29        <valorDefaultAlfanumerico></valorDefaultAlfanumerico>
30      </atributo>
31    </atributos>
32  </metadadoNegocio>
33 </aplicacaoSI>

```

A partir destes dados é possível gerar o código SQL (Código 2) referente à criação do banco de dados com temporalidade, com a tabela de Pessoa Física e os respectivos atributos. A síntese da camada de gerência de temporalidade (que estaria acima da camada de persistência) bem como da aplicação, negócio e interação serão ocultadas por questões de espaço e devem ser apresentadas em um trabalho posterior.

Code 2. Trecho de código gerado a partir do trecho de metamodelo apresentado.

```

1      CREATE DATABASE Aplicacao;
2      CREATE TABLE PesFis(
3          CPF varchar(11),
4          Nome varchar(50),
5          Dt_IniVal TIMESTAMP NOT NULL,
6          Dt_FimVal date,
7          Dt_IniTran TIMESTAMP NOT NULL,
8          Dt_FimTran TIMESTAMP,
9          PRIMARY KEY (CPF, Dt_IniVal, Dt_IniTran)
10     );

```

Como pode-se notar, a tabela agora é temporal. Logo, qualquer alteração em qualquer registro gera uma nova linha na tabela com os dados atualizados. As datas de início de validade e transação precisam fazer parte da chave primária da tabela para permitir que uma pessoa inscrita sob um mesmo CPF possua diferentes registros nesta tabela, cada um correspondente a uma alteração de validade de um dado ou transação realizada sobre aquele dado.

4. Conclusão

Este artigo apresentou decisões de projeto que servem de insumo para a concepção de uma solução dirigida por modelos para a síntese automática de SI temporais.

Como trabalhos futuros destacam-se a possibilidade de expandir a viabilidade desta abordagem para outros SGBD que não apenas o SGBD Postgres; a geração de outras funcionalidades que não apenas as funcionalidades CRUD uma vez que SI modernos demandam outras funcionalidades como a emissão de relatórios, dentre outras; aumentar a flexibilidade de aparência nas telas geradas, uma vez que a utilização do paradigma IF-ELSE engessa as telas geradas nesta metodologia [da Costa 2011, Da Silva and de Oliveira 2009, da Costa and de Oliveira 2010]; tornar o SI gerado orientado a processos.

Além disso, pretende-se conduzir um estudo sobre o impacto da temporalidade no transformador de modelos. Ou seja, a temporalidade deveria ser modelada como um interesse transversal à especificação das regras de transformação, ou deveria ser modular? Ela deve ser especificada apenas no transformador ou o acréscimo ao metamodelo de origem traria algum benefício/malefício, como replicação, problemas de manutenção ou nenhum problema desta natureza?

Cabe ainda investigar, através de testes, a eficiência desta abordagem, isto é, quão grande é o ganho em tempo/recursos humanos ao utilizar esta abordagem ao invés das abordagens tradicionais de desenvolvimento de software.

Como contribuição direta deste trabalho destacam-se as decisões de projeto para a concepção de uma ferramenta dirigida por modelos que permite a síntese de um SI completo e funcional com características temporais, evidenciando as características arquiteturais da solução gerada pela ferramenta e um modo de realizar tal geração.

A implementação destas ideias já está em andamento e já apresenta resultados promissores. Isto evidencia a importância e eficácia da pesquisa e resultados obtidos até o momento, além de ampliar o acervo de projetos em que o Desenvolvimento de Software Dirigido por Modelos pode ser aplicado com sucesso.

References

- Almeida, A. C., Boff, G., and Oliveira, J. L. (2009). A framework for modeling, building and maintaining enterprise information systems software. In *Anais do XXIII Simpósio Brasileiro de Engenharia de Software*, pages 115–125. Fortaleza, Brasil.
- da Costa, S. L. (2011). Uma abordagem baseada em modelos para construção automática de interfaces de usuário para sistemas de informação. Master's thesis, Instituto de Informática - Universidade Federal de Goiás.
- da Costa, S. L. and de Oliveira, J. L. (2010). Construção e manutenção automática de interfaces com o usuário dirigidas por modelos para sistemas de informação. In *Anais do III Workshop de Teses e Dissertações em Sistemas de Informação do VI Simpósio Brasileiro de Sistemas de Informação*.
- da Costa, S. L., Graciano Neto, V. V., Loja, L. F. B., and de Oliveira, J. L. (2010). A metamodel for automatic generation of enterprise information systems. In *Anais do I Congresso Brasileiro de Software: Teoria e Prática - I Workshop Brasileiro de Desenvolvimento de Software Dirigido por Modelos*, volume 8, pages 45–52. UFBA. Salvador, BA, Brasil.

- Da Silva, W. C. and de Oliveira, J. L. (2009). Gerência de interface homem-computador para sistemas de informação empresariais: uma abordagem baseada em modelos. *iSys - Revista Brasileira de Sistemas de Informação*, Vol. 2, 2009.
- de Castro Georg, R. (2010). Bancos de dados temporais: Estudo sistematizado e aplicabilidade em sistemas gerenciadores de bancos de dados atuais. Trabalho de Conclusão de Curso - Instituto de Informática - UFG.
- de Oliveira, J. L., Loja, L. F. B., da Costa, S. L., and Graciano Neto, V. V. (2011). Um componente para gerência de processos de negócio em sistemas de informação. In *Anais do VII Simpósio Brasileiro de Sistemas de Informação*, pages 250 – 261.
- Fowler, M. (2002). *Patterns of Enterprise Application Architecture*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Graciano Neto, V. V. (2012). Evolução de uma arquitetura para frameworks de aplicação de sistemas de informação - uma abordagem de desenvolvimento dirigido por modelos. Master's thesis, Instituto de Informática - Universidade Federal de Goiás.
- Graciano Neto, V. V., da Costa, S. L., and Oliveira, J. L. (2010). Lições Aprendidas sobre Desenvolvimento Dirigido por Modelos a partir da refatoração de uma ferramenta. In *Anais do Encontro Anual de Computação (ENACOMP)*, pages 68–75. Catalão, Brasil.
- Graciano Neto, V. V. and de Oliveira, J. L. (2011). An early aspect for model-driven transformers engineering. In *Proceedings of the 2011 international workshop on Early aspects*, EA '11, pages 7–11, Porto de Galinhas, PE, Brasil. ACM.
- Kleppe, A., Warmer, J., and Bast, W. (2003). *MDA Explained: The Model Driven Architecture - Practice and Promise*.
- Loja, L. F. B., Graciano Neto, V. V., da Costa, S. L., and de Oliveira, J. L. (2010). A business process metamodel for enterprise information systems automatic generation. In *Anais do I Congresso Brasileiro de Software: Teoria e Prática - I Workshop Brasileiro de Desenvolvimento de Software Dirigido por Modelos*, volume 8, pages 37–44, Salvador, BA, Brasil. UFBA.
- Ma, K. and Yang, B. (2010). A Hybrid Model Transformation Approach Based on J2EE Platform. volume 3, pages 161–164, Los Alamitos, CA, USA. IEEE Computer Society.
- Mellor, S. J., Clark, A. N., and Futagami, T. (2003). Guest editors' introduction: Model-driven development. *IEEE Software*, 20(5):14–18.
- Miller, J. and Mukerji, J. (2003). MDA Guide Version 1.0.1. Technical report, Object Management Group (OMG).
- Seidewitz, E. (2003). What models mean. *IEEE Software*, 20(5):26–32.
- Stahl, T., Voelter, M., and Czarnecki, K. (2006). *Model-Driven Software Development: Technology, Engineering, Management*. John Wiley & Sons.
- Wittman, P. (2010). MDA using AndroMDA. <http://www.wittmannclan.de/ptr/cs/mdaandromda.pdf>.

Avaliação de Desempenho de Reconhecimento Automático de Faces Utilizando as Técnicas de Análise de Componentes Principais e de *Fisherfaces*

**Rosana de Paula Lucas, Flávio Henrique Teles Vieira, Weber Brito dos Passos,
Douglas Leo Netzlaff**

Escola de Engenharia Elétrica, Mecânica e de Computação – Universidade Federal de
Goiás (UFG)

Av. Universitária, n. 1488 - Quadra 86 - Bloco A - 3º piso 74605-010 - Setor Leste
Universitário - Goiânia – GO – Brazil

rosana.rpl@hotmail.com, flavio@eeec.ufg.br, weberbrito@gmail.com,
douglas_ntz@hotmail.com

Abstract. *In this paper, we evaluate the face recognition performance of two different techniques based on PCA (Principal Components Analysis) and Fisherfaces. These algorithms are applied to two image banks well-known in the literature, presenting different characteristics. From the comparison of the simulation results, we intend to point out which algorithm provides best face recognition performance for the two image banks under different conditions, for example, by adding noise to the images.*

Resumo. *Neste trabalho, avaliamos o desempenho de reconhecimento de faces em imagens de dois algoritmos diferentes baseados nas técnicas de PCA (Principal Components Analysis) e de Fisherfaces. Estes algoritmos são aplicados a dois bancos de imagens conhecidos na literatura que apresentam características diferentes. A partir da comparação dos resultados de simulação, buscamos apontar qual o melhor algoritmo em termos de desempenho de reconhecimento de faces para as imagens utilizadas sob diferentes condições, como por exemplo, adicionando ruído às imagens.*

1. Introdução

A face é um elemento biométrico individualizado e de fácil obtenção, o que torna a identificação facial uma área de constantes pesquisas.

Os computadores sofrem uma série de limitações e dificuldades para realizar esta tarefa. Para os seres humanos, a identificação e o reconhecimento facial são simples, embora os seus mecanismos ainda não tenham sido totalmente explicados, sendo fonte de grande interesse para os pesquisadores da área de reconhecimento de padrões. Diversos estudos vêm sendo realizados com potencial não só para explicar os mecanismos cerebrais de reconhecimento de faces como, também, para auxiliar no reconhecimento por máquinas/computadores, muito embora tais pesquisas tenham pouca sinergia [Chellappa 1995].

Houve grande crescimento de pesquisas a partir da década de 90, buscando solucionar problemas como encontrar e isolar uma face em uma imagem, e extrair características específicas, como olhos e boca. Várias técnicas foram desenvolvidas e aplicadas, como a Análise de Componentes Principais (PCA – *Principal Components Analysis*), baseada na Transformada de Karhunen-Lòeve (KLT), o *Singular Value Decomposition* (SVD), classificadores baseados em redes neurais, a Análise de Discriminantes Lineares (LDA – *Linear Discriminant Analysis*), dentre outros [Shakhnarovich 2004].

Neste artigo, avaliamos o desempenho de reconhecimento de faces em imagens dos algoritmos PCA e de *Fisherfaces* sob diferentes fatores como número de testes, presença de ruído e variação do número de imagens disponíveis por pessoa, etc.

2. Análise de Componentes Principais

A Análise de Componentes Principais (PCA) foi criada em 1901, por Karl Pearson, matemático britânico, que buscava representar de forma mais simples dados de dimensão elevada. Em 1933, Harold Hotteling usou a técnica de Pearson em matrizes, criando o termo “Componentes Principais”. Em 1987, L. Sirovitch e M. Kirby, no trabalho “*Low-Dimensional Procedure for the Characterization of Human Faces*” [Sirovitch 1987], propuseram o uso do PCA com o objetivo de reduzir as dimensões de imagens, perdendo o mínimo possível de informações. Devido à grande redundância presente nas imagens de face, a técnica teve sucesso nesta área. Após a publicação do trabalho de Sirovitch e Kirby, uma série de outros trabalhos foram desenvolvidos, com destaque para Mathew Turk e Alex Pentland, em 1991 [Turk 1991].

Entretanto, há limitações para o uso do PCA: é necessário haver linearidade no conjunto de dados a serem manipulados, a média e a variância devem ser estatísticas suficientes e, como as maiores variâncias são destacadas, presume-se que elas representam relações significativas, enquanto variâncias menores são menos relevantes.

Apresentamos a seguir os passos do Algoritmo PCA implementado neste trabalho:

Passo 1 Subtração pela média e cálculo da matriz de covariâncias

Primeiramente, deve-se subtrair todos os elementos da imagem pela média da sua dimensão correspondente. Deste modo, a média deste novo conjunto de dados será igual a zero. Calcula-se, então, a matriz de covariâncias. Sendo um conjunto de dados de duas dimensões, a matriz terá a seguinte forma:

$$C = \begin{pmatrix} \text{cov}(x, x) & \text{cov}(x, y) \\ \text{cov}(y, x) & \text{cov}(y, y) \end{pmatrix} \quad (1)$$

Passo 2 Cálculo dos autovetores e autovalores [Turk 1991]

Calcula-se os autovetores e autovalores da matriz de covariância (equação (1)). Neste caso, teremos 2 autovetores. Por terem aparência semelhante a faces, os autovetores são denominados autofaces ou *eigenfaces* conforme mostra a Figura 1.



Figura 1 – Autofaces [Kruger 2004]

Passo 3 Ordenação dos componentes e redução de dimensionalidade

Uma vez calculados os autovetores, pode-se ordená-los em ordem decrescente, de acordo com os autovalores e, conseqüentemente, de importância. Com os autovetores ordenados de acordo com sua importância, podem ser ignorados os autovetores com valores mais baixos. Haverá perda de informação, mas se os autovalores forem pequenos, essa perda será aceitável e, em muitos casos, desejada, visto que muitas vezes estes componentes estarão associados a ruído.

Passo 4 Criação do Vetor de Características e Representação dos Dados

Escolhidos os autovetores que permanecerão no conjunto, deve-se organizá-los em colunas em uma matriz, tal que:

$$\text{VetorCaract} = (\text{autovet}_1 \quad \text{autovet}_2 \quad \dots \quad \text{autovet}_n) \quad (2)$$

Agora, deve-se obter o vetor de características transposto, de tal forma que cada autovetor ocupará uma linha da matriz. O mesmo deve ser feito com a matriz dos dados, subtraídos pelas médias. Assim, cada linha da matriz representará uma dimensão. Fazendo a multiplicação, temos:

$$\text{DadosFinais} = \text{VetorCaractLinha} \times \text{DadosAjustados} \quad (3)$$

sendo *Dadosajustados* os dados iniciais subtraídos da média, com cada linha representando uma dimensão.

Passo 5 Recuperação dos Dados

A recuperação total dos dados só será possível caso nenhum autovetor tenha sido descartado. Quanto maior o número de autovetores ignorados, maior será a diferença entre os dados originais e os recuperados. Reescrevendo a fórmula acima, temos que:

$$\text{DadosAjustados} = \text{VetorCaractLinha}^{-1} \times \text{DadosFinais} \quad (4)$$

Agora, basta somar as médias que foram inicialmente subtraídas para recuperar os dados:

$$\text{Dados Recuperados} = (\text{VetorCaractLinha}^{-1} \times \text{DadosFinais}) + \text{média} \quad (5)$$

Mesmo que autovetores tenham sido descartados, a fórmula acima pode ser usada. Naturalmente, caso todos os autovetores tenham sido mantidos, essa matriz apresentará os dados originais, e caso descartes tenham sido feitos, haverá perda de informação.

3. Fisherfaces

Duas das maiores dificuldades no processo de identificação facial estão relacionadas à variação de iluminação e de expressões faciais. O PCA retém estas informações, que são indesejadas e podem aumentar as taxas de erro, ao extrair as características com maior variância. Dessa forma, os componentes principais obtidos através do PCA podem não ser os mais adequados para discriminação.

Ronald A. Fisher, em 1936, propôs um critério estatístico que visava maximizar a separação entre as classes, e ao mesmo tempo minimizar o espalhamento dentro delas. Esta técnica foi denominada Análise de Discriminantes de Fisher Lineares (FLD – *Fisher's Linear Discriminant*), ou LDA (*Linear Discriminant Analysis*) e é uma técnica supervisionada, visto que necessita-se de ação externa para organizar as classes [Sharmaa 2012].

Em [Belhumeur, 1997], os autores desenvolveram um algoritmo de reconhecimento facial baseado no FLD, com boas taxas de acerto mesmo em situações com variação de iluminação e de expressões faciais. Como as classes de faces estão separadas e nomeadas, pode-se usar estas informações para maximizar a relação entre o determinante da matriz de espalhamento inter-classes com o determinante da matriz de espalhamento intra-classes:

$$\text{Critério de Fisher} = \frac{\det|S_b|}{\det|S_w|} \quad (6)$$

A matriz de espalhamento inter-classes S_b e a matriz intra-classe S_w são definidas por:

$$S_b = \sum_{i=1}^c N_i (\mu_i - \mu)(\mu_i - \mu)^T \quad (7)$$

$$S_w = \sum_{i=1}^c \sum_{x_k \in X_i} (x_k - \mu_i)(x_k - \mu_i)^T \quad (8)$$

sendo μ_i a face média de uma classe X_i e N_i o número de imagens em uma classe. Pode-se encontrar uma base W_{opt} que maximize o critério de Fisher, da seguinte forma [Belhumeur, 1997]:

$$W_{opt} = \arg \max_w \frac{|W^T S_b W|}{|W^T S_w W|} = [w_1 \quad w_2 \quad \dots \quad w_M] \quad (9)$$

em que ϖ_i é o conjunto de autovetores generalizados de S_b e S_w correspondentes aos maiores autovalores.

Neste trabalho, aplicamos o algoritmo proposto em [Belhumeur, 1997], conhecido como técnica de *Fisherfaces* em uma base diferente de imagens considerando também diferentes fatores como adição de ruído e comparando com os resultados do algoritmo PCA.

4. Simulações e Resultados

Buscando uma análise mais ampla, foram utilizados dois bancos de faces com características diferentes: *ORL face database* [Cambridge 2012] e *Yale Face Database* [Yale 2012]. O primeiro banco, *ORL face database*, é composto por imagens que foram geradas entre abril de 1992 e abril de 1994 no *Olivetti Research Laboratory* em Cambridge, Inglaterra. São 40 indivíduos, cada um com 10 imagens distintas, totalizando 400 faces. As imagens foram feitas variando a iluminação, expressões faciais (olhos abertos ou fechados, sorrindo ou não sorrindo), e detalhes faciais (presença ou ausência de óculos). Para alguns dos indivíduos as fotos foram feitas com certo intervalo de tempo. Todas as fotos foram feitas sob fundo escuro e homogêneo. Todas possuem dimensão 112x92 e são em escala de cinza.

O segundo banco de imagens, *Yale Face Database*, é formado por imagens geradas pela Universidade Yale, nos Estados Unidos. Este banco é composto por 165 faces de 15 pessoas. São 11 faces por indivíduo, uma para cada expressão facial, com as seguintes configurações: luz central, feliz, luz à esquerda, com/sem óculos, normal, luz à direita, triste, sonolento, surpreso e piscando. A Tabela 1 apresenta os números de imagens de treinamento e testes utilizadas nas nossas simulações.

Tabela 1 - Relação da quantidade de faces distribuídas por utilidade

Faces					
Origem	Indivíduos	Treinamento		Teste	
		Individual	Total	Individual	Total
Yale	15	10	150	1	15
Cambridge	40	9	360	1	40

Sob condições normais, como mostra a Figura 2, o PCA apresenta uma melhor resposta em termos de taxa de acerto de reconhecimento de face na fase de teste para as imagens de Cambridge, mesmo variando o número de simulações (testes). Contudo, com a inserção de ruído, o PCA perde sua eficácia, enquanto a técnica de *Fisherface* mantém taxas de acerto bem mais satisfatórias, como mostra a Figura 3.

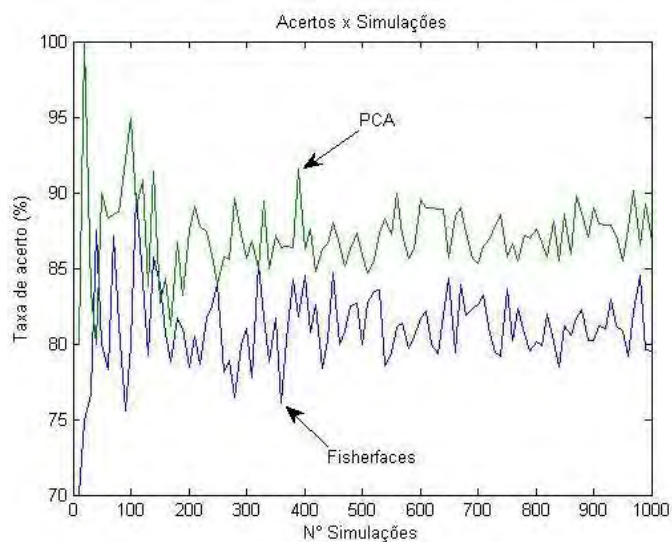


Figura 1 - PCA x Fisherfaces: taxa de acerto versus número de simulações, Cambridge

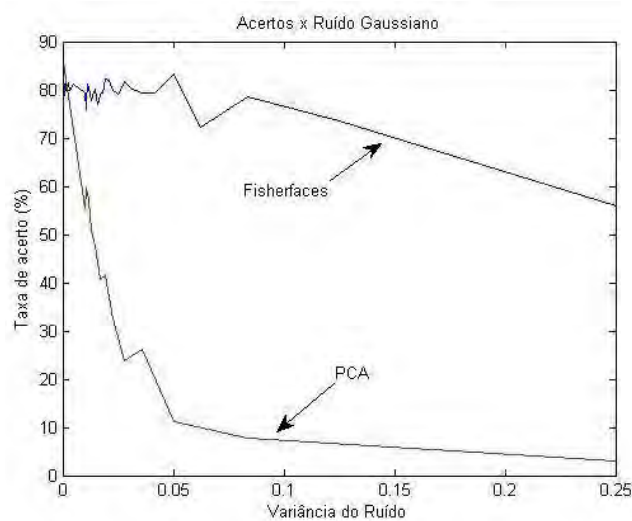


Figura 3 - PCA x Fisherfaces: Adição de ruído gaussiano, Cambridge

Para as faces de Yale percebe-se que a técnica de *Fisherfaces* apresenta desempenho superior em termos de taxa de acerto de reconhecimento de faces, como mostra a Figura 4.

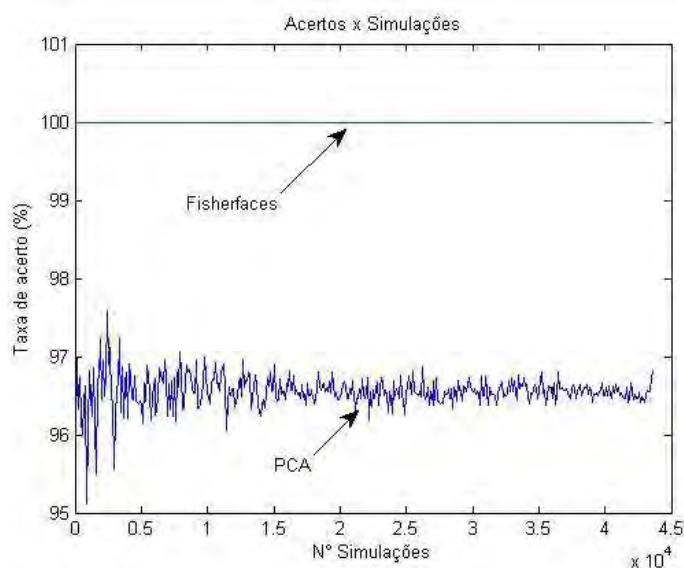


Figura 4 - PCA x Fisherfaces: taxa de acerto versus número de simulações, Yale

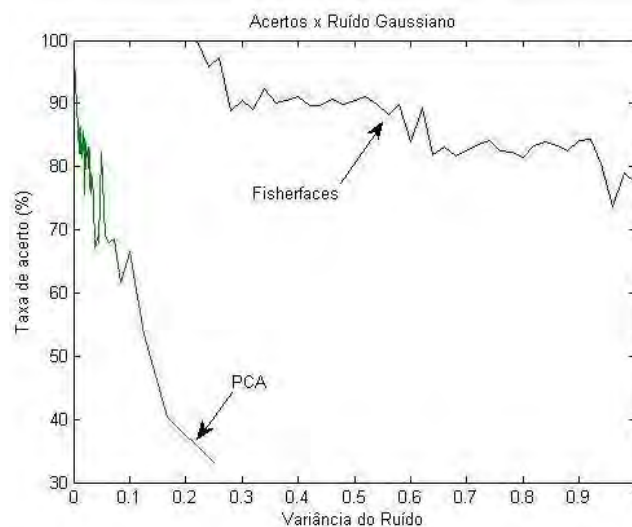


Figura 5 - PCA x Fisherfaces: Adição de ruído gaussiano, Yale

Novamente, observa-se uma maior taxa de acerto de reconhecimento de faces para o algoritmo de *Fisherfaces* quando se aumenta a variância do ruído gaussiano branco nas imagens. Notou-se também que à medida que se aumenta o número de imagens por indivíduo, o algoritmo de *Fisherfaces* obtém melhor resultado do que o algoritmo PCA.

5. Conclusões

Neste trabalho buscou-se simular e analisar dois dos algoritmos utilizados para identificação de faces, a Análise de Componentes Principais e a Análise com *Fisherfaces*.

O método de *Fisherfaces*, que é baseado no FLD, utiliza-se do PCA para possibilitar a extração de características mais discriminantes. As imagens das faces

devem ser agrupadas em classes, cada uma representando uma pessoa. Os dados são organizados de tal forma a maximizar a relação entre o espalhamento entre classes e o espalhamento intra-classes, de forma a facilitar a identificação. Os resultados obtidos nos testes mostram que, de fato, a técnica de *Fisherfaces* obtém melhores taxas de acerto em quase todas as simulações realizadas quando há presença de ruído gaussiano ou variação na iluminação.

Observou-se também variação considerável nos resultados das duas técnicas de acordo com o banco de faces utilizado. Em todos os testes realizados, o banco de Yale possibilitou maiores taxas de acerto. Comparando-se os bancos, Yale possui apenas 15 indivíduos, com 11 imagens para cada um, enquanto Cambridge possui 40, com 10 imagens para cada. Estes aspectos já explicam boa parte das diferenças nos resultados, visto que Yale possui mais informações para discriminar uma quantidade significativamente menor de indivíduos. Deve-se atentar, também, para a qualidade das imagens de cada banco. A resolução das imagens do banco de Yale é 243x320, superior a Cambridge, com resolução de 112 x 92. Yale também privilegiou fotos com variação de expressões faciais, resultando em maiores informações acerca de um indivíduo. A união de todos estes fatores pode explicar as grandes diferenças observadas nas taxas de acerto obtidas em cada banco. Evidencia-se, então, que no desenvolvimento de um sistema de identificação facial deve-se empregar atenção especial na construção do banco de faces, buscando sempre qualidade e o maior número possível de imagens e expressões faciais por pessoa.

Referências

- Cambridge University (2012): Olivetti Research Laboratory. Disponível em: http://www.cl.cam.ac.uk/Research/DTG/attarchive:pub/data/att_faces.zip. Acessado em 10/11/2012
- Chellappa, R. et al. (1995). Human and Machine Recognition of Faces: A Survey, Proc. IEEE, vol. 83, no. 5, p. 705-740.
- Kruger, J. et al. (2004). Obtaining the Eigenface Basis. Disponível em: <http://cnx.org/content/m12531/1.3/>. Acesso em 20 out. 2012.
- Sharmaa, A., Paliwala, K. (2012). A two-stage linear discriminant analysis for face-recognition. Pattern Recognition Letters. Vol. 33, Issue 9, 1 July 2012, p. 1157–1162
- Shakhnarovich, G. et al. (2004). Face Recognition in Subspaces. Handbook of Face Recognition, Eds. Stan Z. Li and Anil K. Jain, Springer-Verlag, Dezembro. 2004
- Sirovitch, L.; Kirby (1987), M. Low-dimensional procedure for the characterization of human faces. Journal of Optical Society of America, v. 4, p. 519-524.
- Turk, M.; Pentland, A (1991). Eigenfaces for Recognition. J. Cognitive Neuroscience, vol. 3, no. 1, 1991.
- Yale Univeristy (2012): Department of Computer Science. Center for Computational Vision and Control. Disponível em: <http://cvc.yale.edu/>. Acessado em 10/11/2012

Análise de Segurança da Camada de Notificação do SparkleShare

Nícolas Kaqui, Marcelo Akira, Bruno Silvestre

¹Instituto de Informática — UFG

{nicolaskaqui, marceloakira, brunoos}@inf.ufg.br

Abstract. *In this work, we analyzed SparkleShare project about some security questions. This project is an alternative for Dropbox or UbuntuOne. We identified some security threats in one component of SparkleShare and we proposed a solution to mitigate the problem.*

Resumo. *Neste trabalho, analisamos quesitos de segurança do projeto SparkleShare, que é uma solução alternativa ao Dropbox ou UbuntuOne para sincronização automatizada de arquivos. Identificamos algumas falhas de segurança em um dos componentes do SparkleShare e propomos uma solução para mitigar o problema.*

1. Introdução

O grande desenvolvimento da sociedade fez com que os serviços básicos e essenciais sejam quase todos entregues de uma forma completamente transparente. As infraestruturas existentes permitem entregar os serviços em qualquer lugar e a qualquer hora, de modo que possamos acender a luz, abrir a torneira ou fazer uma ligação para qualquer lugar. Esse mesmo conceito tem sido aplicado na informática e uma mudança consistente tem sido feita com a disseminação de um novo paradigma, a computação em nuvem.

A computação em nuvem [Taurion 2009, Armbrust et al. 2010] é um paradigma que muda o modelo de computação tradicional, onde todo o processamento de dados e o armazenamento de arquivos é feito em um computador local, para um modelo onde essas tarefas são realizadas por um ou vários computadores na Internet. Esse paradigma pretende ser global e prover serviços para todos, desde o usuário final que hospeda seus documentos pessoais na Internet, até empresas que terceirizam toda a parte de TI.

Com a computação em nuvem, os usuários moverão seus dados e aplicações para a nuvem, podendo acessá-los de forma simples e de qualquer local [Drago et al. 2012]. Sendo uma ferramenta de trabalho importante em ambientes onde ocorre uma troca de arquivos entre vários usuários e para usuários que necessitam acessar seus arquivos de diferentes lugares, como por exemplo, nas empresas, nos escritórios e nas escolas. Os sistemas que permitem a utilização desses documentos em diversos lugares geralmente empregam mecanismos de sincronização para manter a consistência [Xianqiang et al. 2011].

O SparkleShare é um sistema que permite ao usuário manter seus arquivos sincronizados entre várias máquinas. Diferentemente do Dropbox [Dropbox Inc. 2012, Drago et al. 2012] ou UbuntuOne [Canonical Ltd. 2012], cujos serviços são oferecidos por empresas, o SparkleShare é uma solução independente e qualquer instituição ou usuário final podem utilizá-lo para montar seus serviços de sincronização de arquivos.

Em nossos estudos sobre o funcionamento do SparkleShare (versão 0.4.0), identificamos problemas de segurança em seu mecanismo de notificação de sincronização. Neste trabalho, apresentamos a análise feita e como foi realizada a correção do problema.

O trabalho está estruturado da seguinte forma: na seção 2 apresentamos a arquitetura interna do SparkleShare, na seção 3 apresentamos o problema de segurança identificado e a solução adotada, e na seção 4 apresentamos a conclusão do trabalho.

2. SparkleShare

O SparkleShare surgiu como um projeto que permitisse que usuários e empresas pudessem criar o seu próprio repositório de sincronização automatizada de arquivos [Bons 2010].

A possibilidade de trabalhar com mais de um repositório de arquivos e cada repositório estar hospedado em um servidor diferente fez com que o SparkleShare fosse bem visto por empresas e usuários que já possuem um ou mais repositórios e necessitam do sistemas para sincronizar e compartilhar os arquivos [Bons 2012].

O cliente SparkleShare (instalado na máquina do usuário) é composto pelas camadas de sincronização e notificação. A camada de sincronização utiliza um repositório Git para armazenar e manter o histórico dos arquivos, e a camada de notificação utiliza um serviço IRC para troca de mensagens. A Figura 1 apresenta a arquitetura em mais detalhes.



Figura 1. Apresentação das camadas na arquitetura do SparkleShare

A seguir, faremos uma análise das camadas de sincronização e notificação. Apresentaremos o funcionamento e a estrutura de cada uma delas.

2.1. Camada de Sincronização

A camada de sincronização é responsável pelo armazenamento, versionamento e sincronização dos arquivos no repositório central (servidor). O funcionamento dessa camada ocorre em dois momentos: quando o repositório local é modificado ou quando recebe uma notificação remota de alteração do repositório.

Quando o repositório local sofre alguma modificação (geralmente, o usuário modificando seus arquivos), o mecanismo de sincronização registra as modificações no servidor e notifica a mudança pela camada de notificação. Quando o cliente SparkleShare recebe uma notificação de atualização, a camada de sincronização é ativada para que se inicie a sincronização dos arquivos com o repositório central.

A camada de sincronização é implementada utilizando o sistema de controle de versão Git [Loeliger 2009]. Esse sistema foi idealizado e desenvolvido por Linus Torvalds, tendo como principais características a licença livre e o suporte ao desenvolvimento distribuído.

Um conceito importante (e usado na camada de notificação do SparkleShare) é o identificador de *commit* usado pelo Git. Cada versão tem um identificador produzido pela aplicação do algoritmo SHA1 ao conteúdo modificado, produzindo assim um valor de *hash*, por exemplo:

```
68bd599b85538ae08c7729ce8d5127076f567b29
```

O objetivo é que essas referências sejam únicas no repositório e possam ser usadas como referência para o *commit* [Swicegood 2008].

O Git oferece suporte de acesso ao repositório por HTTP, SSH ou protocolo próprio (Git). A versão analisada do SparkleShare utiliza o Git com a comunicação sendo feita pelo SSH, garantindo autenticação do usuário no servidor e canal seguro de comunicação.

2.2. Camada de Notificação

A camada de notificação é utilizada como meio de comunicação entre os clientes SparkleShare. É por essa camada que um cliente avisa aos demais clientes do usuário que é necessário atualizar os arquivos pois foi feita uma modificação no repositório.

A camada é implementada utilizando o protocolo IRC [Charalabidis 1999]. A escolha do protocolo IRC ocorreu pela facilidade de configuração e o baixo *overhead*. Por ser um protocolo de comunicação simples e focado em troca de mensagens simples (apenas texto), o IRC, na época em que foi desenvolvido, se destacou muito pela agilidade de comunicação entre os clientes, mesmo sobre uma conexão considerada hoje lenta [Toyer and Smith 1997].

Dessa forma, o SparkleShare utiliza um servidor IRC como sendo um serviço *Publish-Subscribe*. Os clientes de um usuário se formam uma rede ponto-a-ponto e se conectam ao mesmo servidor IRC para acessar um canal em comum e trocar informações. Assim, o objetivo é que a troca de notificações só ocorra entre eles.

A definição do canal de acesso comum é dado pelo primeiro identificador de *commit* do repositório Git ao qual o cliente faz referência. Podemos ter então canais como a *hash* mostrada na seção 2.1. O uso desse identificador minimiza a chance de colisão no nome dos canais.

Com o objetivo de que fosse fácil para os usuários, em termos de infraestrutura, utilizarem o SparkleShare, o projeto disponibilizou um servidor IRC público de livre acesso, que é usado por padrão. Caso o usuário desejar mudar, basta alterar o arquivo de configuração do SparkleShare.

3. Segurança no SparkleShare

Após a análise mais interna das camadas do SparkleShare, pudemos identificar problemas de segurança quanto ao protocolo utilizado na camada de notificação (IRC).

Por ser de simples implementação, o protocolo IRC não possui uma boa camada de segurança, podendo ser utilizada como meio de possíveis ataques. Por esse motivo, é muito comum ver o protocolo IRC bloqueado por *firewall* em muitos ambientes [Toyer and Smith 1997].

Na construção da camada de notificação do SparkleShare, algumas medidas de segurança do IRC não foram utilizadas, sujeitando o usuário do sistema a alguns possíveis ataques, por exemplo, o ataque de *footprinting* [Assunção 2010]. Nesse tipo de ataque são coletadas informações do usuário, como IP da máquina e canais que ele está utilizando. Usuários que optaram por um servidor IRC privado podem ter o problema reduzido, mas os usuários que fazem uso do servidor público do projeto são mais afetados, devido à exposição.

Utilizando um cliente IRC comum, conectamos no servidor IRC público e realizamos apenas a listagem dos canais existentes, para ilustrar. A figura 2 mostrando a listagem realizada.

```
*** Channel(0) #b621ec88ffef25e6745b19b8fd74c9acal342f(1) #4f4a412b0e796ced1775b4f14ffa5a213912e2f0(1)
*** #df9a4e36cd3f84bd9035e33885848ef84b3e93f(1) #4f8f1431ab97c0db69a2f77a247d068c55cd04b0(1) #53d2f01e20338169b6a9d874cdc28cc36359c7e8(1)
*** #e5f2d86ec5182f1c5f4a46d8b118e9768dc2fa8f(1) #3bc1a349c6c24f01c766624994d32fc9f915806(1) #6a59d1e6332717fd73b49f585f4953211d3130d(1)
*** #7efcbea9e637319e7272f2fae914c093301d(1) #b7b3ae1abfdad5c57bde6f71116814a8c775c377(1) #5a5f98c6271f1bd302db263ca84d7820a8e7bce(1)
*** #9b4b2f5c257648b785fc9bd4d853ae667af1de4(1) #fb3fa4abb150e142a9e8779af18b43e58b91d5ee(1) #ba8832b8c-e3290f070e29f92b6c26b49e4dc61a(1)
*** #7181e569e3a15894f3e5ccc080e13a30da27653b(1) #d935c5ebdc50e663f456a34d723d8bd665f743e2(1) #83d553219eac80ac793e1dea8552624068d2b2097(1)
*** #a3a690aef353b0e0c2850d2d97fb96817f3f5(1) #31ee47e8aee08e59ae13db4d5625f5c f08ac4812(1) #c75717239adfadfae449aaba2896624a98bb0e6(1)
*** #77246e2f7d78d80c14443ca64c9aab5fa31ddb0(1) #348019ae96b054180669b70a07fddc8a05fe2fd99(1) #05129fd9dbfb89080220f6c53ad1335a869c6824(1)
*** #035231bf8ea6493af91de8618b419314637a779(1) #0d3c62e6a6cf03e58a291539afecac9c31b61dbd(1) #110cb87c4bc59d8ea503238e6f6c9ea13f538f8(1)
*** #e2cf28b6e66ed243a93931a4e5327acdf1746(1) #a519bbf8512413a6185e1f85a65b41428aa32c67(1) #c79ff2c8d1c124bd8afed09d5b3d2814a0156cb1(1)
*** #cae7eac847de9f444ed1d6d8e4856bac7d08733f(1) #5179d3027e0bca22536b88f0da8b7d0fd89ba7be(1) #95882641037174022d1eae488242c4b18bc14860(1)
*** #ff03cae5102b534301c025d7ba3dc7be25e26d30(1) #87fbd77383988557b321cfd4600dcaaf29aee91(1) #752af2f013756bba9631d0ff22b5677afde7c8dd(1)
*** #769a76f168598621c3f7a9b4492244bf8fa7b16(2) #267ce8707a66b3f0c6be32dc8567f971363d516(1) #7e02c0691bd81dd3d3d4d24038ac2285168b6e3(1)
*** #8679dad573ba7f8588941b13db13211e874019d6(1) #6b251c32fa647abb40769021b7c3cfaa53c1603(1) #d2b61962a38f3e4f41400c28c1e2d2346c42cc84(1)
*** #0f4f4450892075b6f277deb9334066d11bbde05(1) #360a34be36f7571ed0179e217d8fcf56976726c(1) #9d0ebb3f1a3f88d48c6fb3548625c74c9531b47e(1)
*** #84a1cf2fbccdc3a7040658a081de853ce96ee68(1) #1e372e0d0c181e12185af33a6ffebc36fd598efc(1) #e88cb511fb31a87919eaceb7af4d4e1d0a45a05e(1)
*** #014caed24b35dd53f2f0a5fbb02101390c6d50d(1) #2bf1599bc9ce0fb1bdf20d7f3bb4813ded6c5aa1(1) #d071e4b4b9f53527879e546bcb871a9c594de29f(1)
*** #d59e6f40532913ee0948b3194417303f4ce37d7(1) #97b38139b61b0822a185c32b158ea68d2a547484(1) #e7cd864b557f2a0b4b7fdae77f68552b4571(1)
*** #2df9b42483abb028d4a0ab57accab829dc212d5(1) #0a610db7cde3b57ba8ad30a8e6b57dfb709d60e5(1) #e5757a31c220bfd6330a9627048dc385afbc0a9(1)
*** #2e2ccdf4d2eb4f5223963dd6aae10db08c2d6481(1) #107928ff3e00089f5cb35932cb8c48824d3d90fd(1) #c0865be79705df133ae6f6c4986c6a5d59bb913(1)
*** #153a4cf36d3b1193e2219fdd206e5ae5ebda5a87(1) #df95bfc4bbe9e663bfff36498e5f6b15028675fa5(1) #51a48d9421fef3a1eb890529be63b4dc53ac74(1)
*** #3a6364523b3a272c44a2d08e36a6e1681ada5a6(1) #e0b06a145d6f4637cf90238592f3d3c290ba4376d(1) #12ef7b51cf29594c1a4e7134b15134e05f62e672(1)
*** #171ba8c3f444434ae12ab50721b3dd7a3a165db(1) #077a2e9fa16bcaaba10e7101a438eb5aa74f2804(1) #c5cd564cf707f0ee97250566a47c5259ae4b1d439(1)
*** #04dbba29b306a90ca108ec8093dab0991fa68e93(1) #d5a8c6bc8e8708d2bf527480cf048eeb601a2b1(1) #7952e4a586bac5daa0c5424bf9b9bd511ad45717(1)
*** #d0a4e558db6a1940bc39beb89a28560716d3a0e(1) #e7c9bfc89d1b5094a315ababdff659a46ddcc4bc3(1) #e3917f653a85115e7cbf439926fc4bf0e56558b5(1)
*** #6ba4e0dc0b43954acbl1ae675311e5f68ac67b3(1) #f864c1c8e31a339c599abab3b5ca543322b6e1d(1) #6176508304c40d4a2f44b0737c0e6d3bb7b3c4(1)
*** #10e5fa156d1478dd2a5e8246182c959bd16e71d(1) #ce218f1c f889fd731dc11649846ef37a0ebfb3b(1) #7541f21a554bf91558d2c77147e0e83ae310ac9(1)
*** #908a4972128f438c24d398395a0ccdcfa11539b(3) #7ae91ab4770839a56d3a3c0f6b30180f019fa55f(2) #7c6324360cfcc6e3f09398b69408547fb3ae1(1)
*** #3a2daaaf1728fed9df1fb95c2ae4575d22f8b0ce(1) #441184bd8b901b5979a553c49d2ede7756c247a(1) #17c96edf4c492c3aa7d3db6b0c8fa17ae2f57(1)
*** #49b4d46dc67a2a05df7736877501819e4d4e75a0(1) #79aa65d8c872dc56c5f73313bda7ef204db108(1) #d39ed54d1cb4d63227c5b30b13a40475f7fb2c7ec(1)
*** #204d2f1257fd375b98bb411d966098fcl17ae5f0(1) #1017289523056eaa0d1c1d2e82a30431f83bd6b1(1) #dc722e004a96d8d8d307b2f475d219e357152e5c(1)
*** #45bacac495c5e76312fd311fad4c607546924f(1) #1cb8b78eb13895373fa7cb3aa91f5997164bbe2(1) #229b16536bd837a691a1cda0c6983453539986(1)
*** #43cbcf8db8151028fafecba6c5c6875a1e59e8e6d(1) #5fb9c615ba5554ba82362047270f726b69db847a(1) #59dec12ad34c5f1735d94d27c087f32158838b4(1)
*** #c56a4f7470f941361c667ae82adca2a877f083d3(1) #d871481ecdd7f1bdc89e7d6ec7d9e137b5b5(1) #67e9c9d85687f68a9c6046db1d3df2bb4ae0(1)
*** #1cb70d4c0331128951a2fcc9a40bcf9a06113da(1) #af1b749131ed75b2f42340da8d6ec4db29290b5b(1) #3fcfd92e8bc2f4f4a4ec9f250967e14d3c72c6c0f(1)
*** #b680124037c4e4ae2d4849b3e9985b04dbec32c(1) #3fd496e8da80362b5f1807fda4b443a2ce7c2e2d(1) #205aaa0cfd99497f76d0ec80e4a751a034764fe4(1)
*** #294af7f40d53037bc8243a32d49aa0dbd0ed59(1)
```

```
[1] 16:03 on #908a49727 (+nt) * type /help for help
```

Figura 2. Listagem dos canais no servidor IRC do SparkleShare.

Além disso, os canais criados pelo SparkleShare não possuem nenhuma proteção. Para exemplificar, como mostrado na Figura 3, utilizando novamente um cliente IRC comum, conseguimos acessar o canal de nossa instalação do SparkleShare. Vários autores citam que a segurança de sistemas baseados em publish/-subscribe ou P2P começa com a garantia de que não há vazamento ou infiltração de mensagens externas [Bacon et al. 2008, Singh et al. 2011, Bailes and Templeton 2004, Zhang et al. 2005]. Pode-se notar que essa premissa foi claramente violada na implementação do SparkleShare que analisamos.

Como o cliente SparkleShare não possui nenhum formato para a mensagem de

```

*** [redacted] has joined channel #908a49727128f438c24d3983958a0ccdcacf1d39b
*** #908a49727128f438c24d3983958a0ccdcacf1d39b 1322568746

[1] 16:02 on #908a49727 (+nt) * type /help for help
/join #908a49727128f438c24d3983958a0ccdcacf1d39b

```

Figura 3. Acessando um canal IRC do SparkleShare por cliente IRC externo

notificação, qualquer mensagem postada no canal (até mesmo uma simples letra) é interpretada como uma indicação de que o repositório foi alterado. Somando isso ao acesso aberto aos canais, é possível enviar falsas notificações.

Nesse caso, é possível construir um pequeno programa que varre todos os canais do IRC enviando uma mensagem qualquer. Isso fará todos os clientes irem ao repositório central, ao mesmo tempo e repetidamente, para realizar a atualização da base local. Esse volume de requisições pode causar desde um transtorno até tornar o servidor inacessível — causando o chamado *Denial of Service* [Assunção 2010].

3.1. Soluções de Segurança

Tendo identificado algumas vulnerabilidades do SparkleShare 0.4.0, foram estudadas algumas soluções de segurança que tornam a camada de notificação mais segura. As soluções propostas para garantir a segurança dos usuários e dos canais dentro do IRC são: canais no modo secreto e canais com senha.

A especificação do IRC permite que o usuário configure um canal em modo secreto. Esse tipo de canal não poderá ser listado pelos usuários, dessa maneira é necessário que o usuário do canal saiba o nome do canal para poder acessá-lo [Charalabidis 1999].

Essa solução contribui para reduzir as informações disponíveis, e como o *hash* é utilizado como nome do canal, isso oferece um nível maior de segurança. Por outro lado, não afeta o funcionamento dos clientes, dado que eles possuem o identificador do repositório Git.

Entretanto, a abordagem descrita não protege do fato do *hash* ser conhecido de antemão (de acessos inseguros anteriores) ou gerado (pois é apenas uma string). E como soluções de segurança não podem ser simplesmente baseadas na ocultação de informações, a segunda medida que propomos é o uso de senha para acesso ao canal oferecido pelo protocolo IRC.

Definimos que os clientes SparkleShare ligados a um mesmo repositório Git sejam configurados com uma senha. O primeiro cliente a se conectar com o servidor IRC cria o canal especificando a senha compartilhada, os demais necessitam da senha para fazer parte do canal.

Mas temos que considerar ainda que o canal pode já existir e ele não estar protegido por senha, o que leva a dois casos: (i) um cliente sem suporte à senha foi ativado primeiro ou (ii) alguém criou o canal para esperar conexões de clientes e aplicar algum ataque.

Considerando a possibilidade do caso (ii), a melhor abordagem seria impedir que o cliente se conectasse com um canal não protegido. Mas, os clientes SparkleShare de um usuário podem estar instalados em diversas máquinas e o processo de atualização não é feito de forma instantânea. Dessa forma, os novos clientes atualizados para a nossa versão não poderiam funcionar até todos os clientes estiverem atualizados — o que pode gerar um transtorno para o usuário.

Para se acomodar a esse período de transição, optamos então por uma solução em que o usuário pode escolher se o cliente deve ou não se conectar a um canal não protegido por senha. Dessa forma, o usuário poderá atualizar os clientes sem necessariamente impedi-los de funcionar. Quando todos os clientes estiverem atualizados, o usuário poderá então ativar a proteção.

Novos usuários não terão o problema de atualização e poderão iniciar já com a proteção de canais ativada.

3.2. Comentários sobre a Implementação das Soluções

Durante o processo de aplicação das modificações no código fonte do SparkleShare, nos deparamos com o fato da solução dos canais em modo secreto já ter sido implementada na versão de desenvolvimento [Bons 2011]. Focamos então na implementação dos canais com senha.

O SparkleShare é desenvolvido em C# e possui uma classe (`SparkleListenerIrc.cs`) que realiza todo o processo da comunicação do sistema com o cliente IRC. Essa classe por sua vez utiliza a biblioteca `SmartIrc4net` [Bauer 2012]. Modificamos então o trecho de conexão no canal, introduzindo o modo protegido por senha.

A configuração da senha e das políticas de acesso (forçar o acesso somente por senha ou não) aos canais é feita no arquivo de configuração XML padrão do SparkleShare. No sistemas operacional Debian, onde os testes foram realizados, o arquivo de configuração é o seguinte:

```
/home/user/.config/sparkleshare/config.xml
```

O usuário deve incluir as seguintes propriedades nesse arquivo de configuração para que a proteção funcione adequadamente:

- **announcements_password**: senha do canal.
- **allow_passwordless_join**: booleano que indica a permissão para acessar canais não protegidos.

As modificações foram testadas com um repositório no GitHub [GitHub Inc. 2011] e com o servidor público IRC do SparkleShare. Depois dos testes, um *patch* foi oferecido aos desenvolvedores do SparkleShare, que incorporaram as nossas modificações no repositório do projeto [Kaqui 2011]

4. Considerações Finais

Com a evolução da computação em nuvem e a popularização dos serviços oferecidos por ela, foram desenvolvidos os sistemas de sincronização e compartilhamento de arquivos com a finalidade de abstrair a utilização do armazenamento de arquivos em nuvem e de oferecer um serviço de sincronização de arquivos locais com a nuvem. O SparkleShare é um sistema de sincronização desenvolvido dentro da filosofia de Software Livre e permite que os usuários e instituições utilizem diferentes repositórios de armazenamento.

Por utilizar elementos disponíveis na Internet, o funcionamento do SparkleShare está sujeito a ataques. Neste trabalho realizamos uma análise de segurança do SparkleShare, versão 0.4.0, identificamos alguns pontos falhos em seu serviço de notificação, e aplicações soluções que aumentam a segurança.

Os desenvolvedores do sistema já reconheceram que o uso do IRC não é uma boa opção, sendo que na época em que este trabalho foi realizado, o desejo dos desenvolvedores era substituir pelo protocolo XMPP [Foundation 2004]. Durante a escrita deste artigo, verificamos que o projeto agora utiliza uma solução própria para notificação.

Ainda assim, acreditamos que a nossa solução de canais protegidos por senha implementada era de importância para o projeto, pois oferecia um nível melhor de segurança para os usuários, mantendo compatibilidade com versões anteriores e de baixo impacto no código (comparado com adaptar a camada de notificação para um novo protocolo).

Destacamos novamente que essa melhoria foi enviada aos desenvolvedores do SparkleShare e incluída no código final do sistema.

Referências

- Armbrust, M., Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., and Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53:50–58.
- Assunção, M. (2010). *Guia do Hacker Brasileiro*. VisualBooks.
- Bacon, J., Eysers, D. M., Singh, J., and Pietzuch, P. R. (2008). Access control in publish/subscribe systems. In *Proceedings of the Second International Conference on Distributed Event-based Systems*, DEBS '08, pages 23–34, New York, NY, USA. ACM.
- Bailes, J. E. and Templeton, G. F. (2004). Managing P2P security. *Commun. ACM*, 47(9):95–98.
- Bauer, M. (2012). SmartIrc4net - C# IRC Library. <http://www.meebey.net/projects/smartirc4net/>.
- Bons, H. (2010). Announcing SparkleShare. Design Monkey. <http://www.bomahy.nl/hylke/blog/announcing-sparkleshare/>.
- Bons, H. (2011). GitHub commit description. <https://github.com/hbons/SparkleShare/commit/5f31d5f43d41a34300eee886bef9bc8e1a52800c>.
- Bons, H. (2012). SparkleShare wiki. <https://github.com/hbons/SparkleShare/wiki>.

- Canonical Ltd. (2012). Ubuntu One. <http://one.ubuntu.com>.
- Charalabidis, A. (1999). *Book of IRC: The Ultimate Guide to Internet Relay Chat*. The Book Of Series. No Starch Press.
- Drago, I., Mellia, M., M. Munafo, M., Sperotto, A., Sadre, R., and Pras, A. (2012). Inside Dropbox: Understanding personal cloud storage services. In *Proceedings of the 2012 ACM Conference on Internet Measurement Conference, IMC '12*, pages 481–494, New York, NY, USA. ACM.
- Dropbox Inc. (2012). Dropbox. <https://www.dropbox.com/>.
- Foundation, J. S. (2004). Extensible messaging and presence protocol (XMPP): Core. RFC 3920.
- GitHub Inc. (2011). GitHub. <https://github.com/>.
- Kaqui, N. (2011). GitHub commit description. <https://github.com/hbons/SparkleShare/commits/0.4.0/SparkleLib/SparkleListenerIrc.cs?author=nicolaslazartekaqui>.
- Loeliger, J. (2009). *Version control with Git*. O'Reilly Series. O'Reilly.
- Singh, J., Eysers, D. M., and Bacon, J. (2011). Disclosure control in multi-domain publish/subscribe systems. In *Proceedings of the 5th ACM International Conference on Distributed event-based system, DEBS '11*, pages 159–170, New York, NY, USA. ACM.
- Swicegood, T. (2008). *Pragmatic Version Control Using Git*. Number v. 1 in Pragmatic Version Control Using Git. Pragmatic Bookshelf.
- Taurion, C. (2009). *Cloud Computing - Computação em Nuvem: Transformando o Mundo da Tecnologia da Informação*. BRASPORT.
- Toyer, K. and Smith, N. (1997). *Learn advanced Internet relay chat*. Wordware Pub.
- Xianqiang, B., Nong, X., Weisong, S., Fang, L., Huajian, M., and Hang, Z. (2011). Sync-views: Toward consistent user views in cloud-based file synchronization services. In *Chinagrid Conference (ChinaGrid), 2011 Sixth Annual*, pages 89–96.
- Zhang, X., Chen, S., and Sandhu, R. (2005). Enhancing data authenticity and integrity in P2P systems. *Internet Computing, IEEE*, 9(6):42–49.

Estimação do Tempo Médio de Resposta do Servidor em Plataformas de Ensino a Distância do Tipo Cliente-Servidor Utilizando Modelo Baseado em Teoria de Filas

Victor Hugo Teles Costa, Jonatas Abreu Fernandes, Bruno Henrique Pereira Gonçalves, Flávio Henrique Teles Vieira, Leonardo Antônio Alves

Escola de Engenharia Elétrica, Mecânica e de Computação – Universidade Federal de Goiás (UFG)

Av. Universitária, n. 1488 - Quadra 86 - Bloco A - 3º piso 74605-010 - Setor Leste
Universitário - Goiânia – GO – Brazil

victor@vcosta.com.br, jonatasfdv@gmail.com, brunohpg@gmail.com
flavio@eee.ufg.br, leoaalvs@gmail.com.

Abstract. *In this paper, we present an analytical model to predict the performance of Client-Server based e-Learning Platforms. Our approach is based on the queueing theory and can adequately describe some performance parameters of real e-Learning Platforms such as server response time. We evaluate the modeling capability of the proposed approach by comparing the results provided through the model to the results obtained by other model presented in the literature. However, our approach also takes account of the data loss probability in the system. Most importantly, in this study, we develop a tool for designing and scaling Client-Server based e-Learning Platforms based in client-server architecture.*

Resumo. *Neste trabalho, apresentamos um modelo para prever o desempenho de plataforma de Ensino a Distância baseadas em cliente-servidor. A nossa abordagem é baseada em Teoria de Filas, focando principalmente na estimação do tempo de resposta do servidor. Avaliamos a eficiência do modelo proposto para a plataforma de Ensino a Distância, comparando os resultados fornecidos deste com os de outro modelo apresentado na literatura. Entretanto, a principal diferença é que nosso modelo leva em consideração a probabilidade de perda de dados. Assim, neste trabalho, desenvolvemos uma ferramenta para a concepção e dimensionamento de plataformas de Ensino a Distância baseadas na arquitetura cliente-servidor.*

1. Introdução

Desde o final do século XX, a TI (Tecnologia da Informação) tem oferecido diferentes paradigmas para o desenvolvimento da Tecnologia Educacional. Hoje, na era da informação, apesar de suas dificuldades e restrições, é importante que as pessoas tenham acesso à educação formal. Neste contexto, o Ensino a Distância (EaD) está emergindo como um novo paradigma, permitindo que a informação seja compartilhada entre todos os públicos, removendo as barreiras físicas e sociais.

Neste novo cenário, os educadores podem criar laboratórios e salas de aula virtuais onde os alunos contam com materiais de ensino de alta qualidade através de ambientes multimídia, incluindo ilustrações tridimensionais, armazenamento de arquivos, videoconferências e reuniões online.

As empresas e instituições que pretendem explorar o Ensino a Distância precisam construir uma plataforma que deve ser suportada por uma infraestrutura de telecomunicações composta principalmente por servidores de rede, permitindo o uso de todo tipo de mídia baseada em computador.

O Ensino a Distância em todo o Brasil e no exterior é basicamente realizado através de sites, um conjunto de páginas, estáticas ou dinâmicas, organizadas numa estrutura conhecida como WWW (*World Wide Web*). Portanto, a *web* tem um papel crucial no mundo da informação, sustentando as chamadas Plataformas de EaD.

O desempenho da Plataforma de EaD do tipo Cliente-Servidor depende diretamente do tempo de resposta do servidor nas conexões, isto é, o atraso para responder à requisições de página [Drakopoulos e Merges 1992] [King 2003]. Por esta razão, focamos no tempo de resposta e nos parâmetros de rede envolvidos na sua estimação. Para calcular o desempenho da Plataforma de EaD, examinamos um modelo Cliente-Servidor baseado na Teoria de Filas [Gross, Shortle, Thompson, e Harris 2009] [Kleinrock 1976][Pujolle e Gelenbe 1998]. Verificamos que o modelo de fila proposto pode ser utilizado para auxiliar no projeto de Plataformas de EaD, assim como para verificar sua escalabilidade.

O restante deste trabalho é organizado da seguinte forma: na seção II, apresentamos uma abordagem para modelagem de Plataformas de EaD baseada na Teoria de Filas, na seção III, avaliamos o desempenho de Plataformas EaD em termos do tempo médio de resposta do servidor de acordo com a equação proposta, finalmente, na seção IV, relatamos nossas conclusões.

2. Tempo Médio de Resposta do Servidor em Plataformas de EaD

Nesta seção, propomos um modelo para estimar o tempo médio de resposta do servidor em plataformas de EaD baseadas na arquitetura cliente-servidor.

No modelo proposto, consideramos que há três fases no processo de atendimento das requisições de usuários na plataforma de EaD: a chegada de uma solicitação, o processamento no servidor e a chegada do processamento no cliente.

A Figura 1 mostra a rede de filas proposta utilizada para a obtenção das equações para o cálculo do tempo de resposta do servidor, onde A é o número de requisições de chegada ao servidor, S é o tamanho da banda (taxa de transmissão) do servidor e C , o tamanho da banda (taxa de transmissão) do cliente e p é a probabilidade ou taxa de perda de pacotes.

Denominamos de S_r , S_s , S_c , as capacidades de atendimento das filas referentes ao servidor Web, à conexão Internet e ao *browser* do usuário, respectivamente. Nesta rede de filas, as requisições de entrada A se juntam na entrada no servidor de entrada de requisições S_r às requisições A' que deverão ser retransmitidas.

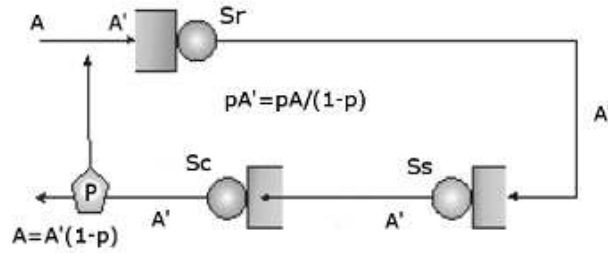


Figura 1. Modelo para Plataforma de EaD.

Essas requisições entram no servidor S_s para serem processadas, e são entregues ao cliente S_c , que poderá solicitar retransmissões ao servidor Web S_r . Denotaremos por C a banda, ou taxa de transmissão do usuário S_s (conexão Internet) e por S a banda do servidor, ou taxa com a qual o servidor responde à Internet. É fácil demonstrar que a capacidade S_r do servidor Web é dada por:

$$S_r = \frac{1}{\left(y + \frac{B}{R}\right)} \quad (1)$$

onde F é o tamanho do arquivo recebido, B é o tamanho do *buffer*, y é o tempo de espera no servidor, R é a taxa dinâmica de atendimento, ou taxa em que o servidor processa o *buffer*.

Podemos enunciar a seguinte proposição referente ao tempo médio de resposta para plataformas de EaD.

Proposição 1: Assumindo que o tráfego em redes de EaD seja Poissoniano, o Teorema de Jackson é aplicável ao modelo de filas apresentado na Figura 1. Assim, o tempo de resposta médio de um servidor EaD utilizando arquitetura cliente-servidor é dado por:

$$E\{T\} = \frac{1}{\frac{1}{y + \frac{B}{R}} - \frac{A}{(1-p)}} + \frac{1}{\frac{C}{F} - \frac{A}{(1-p)}} + \frac{1}{\frac{S}{F} - \frac{A}{(1-p)}} \quad (9)$$

De acordo com Alves, Vieira e Araújo (2009), o modelo de filas que descreve uma Plataforma EaD pode ser considerada uma rede de Jackson e, portanto, o tempo de espera poderia ser calculado através de:

$$E\{T\} = \frac{F}{C} + \frac{I}{1 - AI} + \frac{F}{S - AF} + \frac{F(B + Ry)}{BR - AF(B + Ry)} \quad (10)$$

onde I corresponde ao tempo de inicialização.

Para validar a equação proposta de tempo médio de resposta do servidor em plataformas de EaD, inicialmente implementamos um simulação de fila simples com um servidor com *buffer*. A simulação foi realizada no *software* Rockwell Arena 12, versão estudantil, considerando um *buffer* de 50 Mbits de espaço, uma banda de servidor de 30 Mbps e taxa de requisições Poissonianas.

Na próxima seção, apresentaremos uma comparação dos resultados obtidos com as Equações 9 e 10.

3. Estudos de Otimização de Desempenho de Plataforma de EaD

Analizamos o desempenho da plataforma de EaD em termos do tempo de resposta do servidor, variando parâmetros como a banda do servidor S e o tamanho do arquivo F . Para tal, consideramos como parâmetros para o modelo, valores reais obtidos em uma plataforma real de Ensino a Distância mantida pela empresa NT educação no SENAI-GO.

Calculando os tempos de resposta utilizando a equação proposta temos que, após certo valor de banda (capacidade do servidor), o tempo de resposta do servidor diminui insignificativamente. O que significa que a contratação de mais banda não corresponde necessariamente a uma diminuição no tempo de resposta do servidor. A variação no tempo de resposta foi de menos de 5%. Isso implica que uma plataforma de EaD não depende estritamente da quantidade de banda, mas também da capacidade de atendimento às requisições. A Figura 2 corrobora essa afirmação, onde fizemos $p=0$ e fixamos os valores de $B=2000$ bytes, $R=25547$ bytes/s, $C=38400$ bytes/s (ou 300 kbits/s), $A=3,715$ requisições/s, $F=1500$ bytes, $y=1$ ms e $I=0$ s e variamos a capacidade S do servidor. Note que os valores considerados para B , R , C , A , y e I são valores que foram encontrados em uma plataforma real de Ensino a Distância mantida pela empresa NT educação no SENAI-GO.

Observamos que para os parâmetros de rede considerados, a variação da capacidade do servidor S a partir de 185855 bytes/s não produz variação significativa no tempo de resposta (variação menor do que 5%). Pode-se então, afirmar que o tempo de resposta independe da capacidade para $S > 185855$ bytes/s, no cenário considerado.

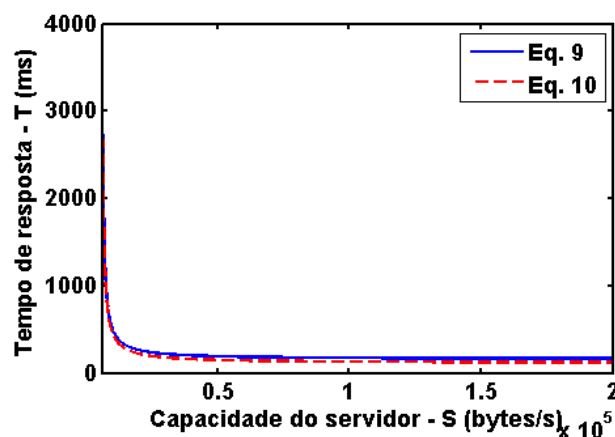


Figura 2. Tempo Médio de Resposta versus Capacidade do Servidor (S).

Quanto maior o tamanho do arquivo (F), maior o tempo de resposta, conforme apresenta a Figura 3 (para o caso de $p=0$). A Figura 3 foi gerada considerando os valores dos parâmetros do sistema fornecidos anteriormente, fixando S em 393216 bytes/s (ou 3 Mbits/s) e variando F .

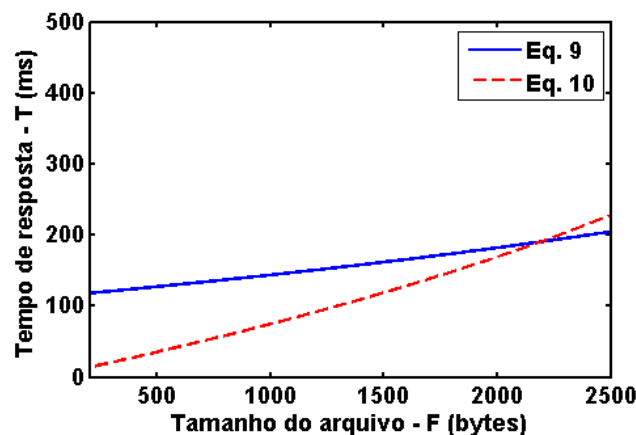


Figura 3. Tempo de Resposta versus Tamanho do Arquivo (F).

A otimização da Plataforma de EaD, com respeito ao tempo de resposta, deve levar em conta a relação entre a variável $E\{T\}$ e as variáveis C , B , S e R . Podemos ajustar os parâmetros C , S e R a fim de obtermos uma melhoria do desempenho do sistema, considerando a disponibilidade de cada recurso. Por exemplo, para obtermos uma redução mais significativa do tempo de resposta no sistema caracterizado pela Figura 3, com $S=185855$ bytes/s, temos de aumentar os parâmetros S ou R dentro dos limites disponíveis.

Ainda considerando $B=2000$ bytes, $R=25547$ bytes/s, $S=393216$ bytes/s, $A=3,715$ requisições/s, $F=1500$ bytes, $y=1$ ms, $I=0$ e $p=0$, na Figura 4, observamos a relação que se desenvolve entre o tempo de resposta e a capacidade C do cliente. Claramente, quanto maior a capacidade C , menor será o tempo de resposta do servidor. No entanto, há um momento em que um aumento na banda pode não significar uma diminuição considerável no tempo de resposta, dependendo da disponibilidade de recursos. Mais especificamente, nas circunstâncias da Figura 4, podemos dizer que para $C > 250709$ bytes/s a variação no tempo de resposta é de menos de 5%, para $S=393216$ bytes/s.

Para os parâmetros considerados anteriormente ($p=0$), e fixando $C=38400$ bytes/s, o comportamento do tempo de resposta em relação à taxa R de atendimento pode ser observado na Figura 5.

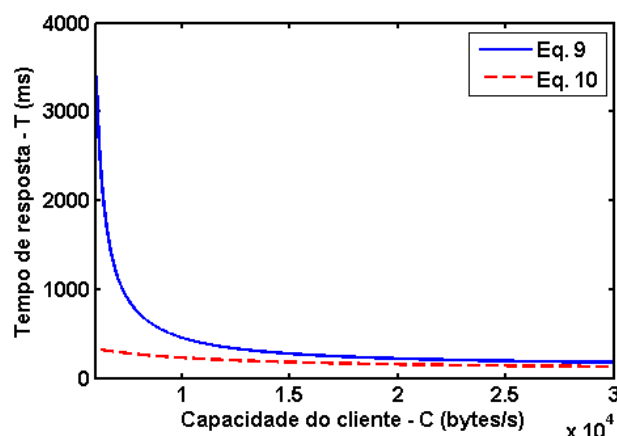


Figura 4. Tempo de Resposta versus Capacidade do Cliente (C).

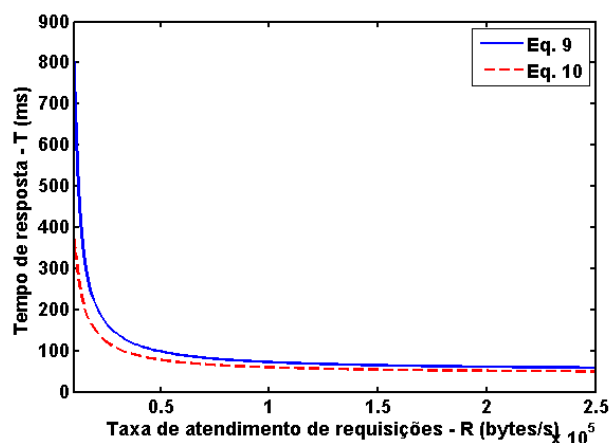


Figura 5. Tempo de Resposta versus Taxa de Atendimento de Requisições (R).

Como na relação entre tempo de resposta e capacidade do cliente, notamos uma diminuição no tempo de resposta com o aumento da capacidade de processamento de requisições. No caso da Figura 5, observamos que, para $R > 764556$ bytes/s, o tempo de resposta não tem diminuição significativa para $p=0$, a variação é inferior a 5%.

Atribuindo alguns valores para p , para que se possa analisar a influência das perdas no sistema, tem-se os resultados de tempo médio de resposta apresentados na Tabela 1. O valor de p corresponde à porcentagem de pacotes que não chegam ao destino devido a erro no pacote ou falha no sistema, sendo que devem ser retransmitidos pelos mecanismos do TCP-IP. Note que nos arquivos de log da plataforma de Ensino a Distância mantida pela empresa NT educação no SENAI-GO encontra-se perdas próximas a 0,1%.

Analizando os resultados de tempo médio de resposta para diferentes valores de p , constatamos que a variação do tempo de resposta $E\{T\}$ em relação a p diferente de zero é pequena. Para $p=0$, notamos uma relativa aproximação do tempo de resposta dado pela equação proposta com o tempo de resposta da Equação 10, como mostra a Tabela 1.

Tabela 1. Tempo de Resposta (em milissegundos) considerando diferentes valores para p

A	E{T} (eq. 9)			E{T} (eq.10)
	$p=0$	$p=0,001$	$p=0,01$	
2,97222	151,7845	151,8224	152,1677	115,1528
3,34375	156,6858	156,7316	157,1492	117,1502
3,71528	161,9595	162,0143	162,5146	119,2606
4,08681	167,6521	167,7174	168,3128	121,4939
4,45833	173,8189	173,8961	174,6016	123,8612
4,82986	180,5252	180,6163	181,4498	126,3748
5,20139	187,8493	187,9568	188,9401	129,0490
5,57292	195,8859	196,0126	197,1726	131,8995
5,94444	204,7505	204,9000	206,2701	134,9444
6,31597	214,5856	214,7623	216,3844	138,2044
6,68750	225,5690	225,7787	227,7060	141,7032

A Figura 6 revela uma influência significativa da perda p no tempo de resposta ao variar a capacidade do cliente (C). Observamos um incremento claro do tempo de resposta com o aumento no valor de p . Como era de se esperar, quanto menor o valor de p , a curva de tempo de resposta se aproxima mais da curva para $p=0$. Para $p=0,1\%$, como observado na Tabela 1, os valores são praticamente os mesmos para $p=0$.

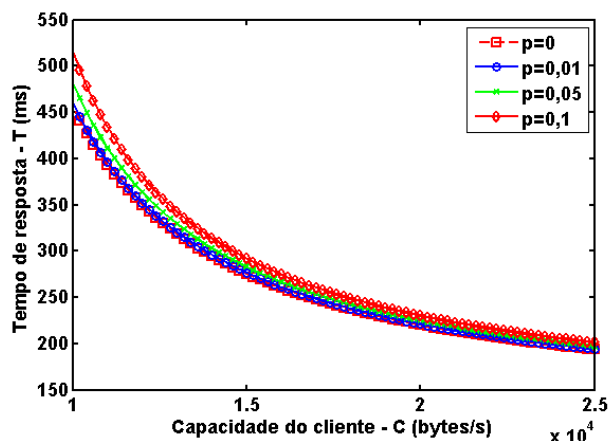


Figura 6. Tempo de Resposta versus Capacidade do Cliente (C).

Na Figura 7, observamos que ao alterar o tamanho do arquivo transmitido, há uma diferenciação considerável dos valores de tempo de resposta para diferentes valores de p quando o arquivo apresenta tamanhos superiores a 4000 bytes. Ou seja, para valores de tamanho de arquivo abaixo de 4000 bytes, o impacto da taxa de perda de dados no tempo de resposta é menor.

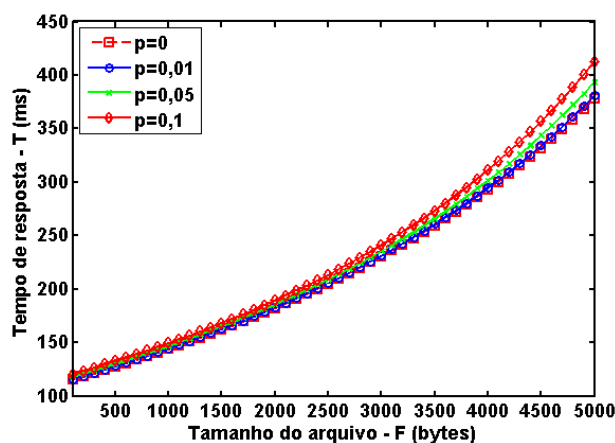


Figura 7. Tempo de Resposta versus Tamanho do Arquivo (F).

4. Conclusões

Neste trabalho, apresentamos um modelo de Plataforma de EaD do tipo Cliente-Servidor baseado em Teoria de Filas. Comparamos o Tempo de Resposta do servidor dado pelo nosso modelo ao obtido com o trabalho desenvolvido em Alves, Vieira e Araújo (2009), verificando a consistência dos valores estimados com o modelo desenvolvido.

Na equação proposta, levamos em conta a influência da taxa de perda de pacotes. Através disso, pudemos observar como a perda pode se relacionar ao aumento do tempo de resposta de uma Plataforma EaD. Também avaliamos a relação do tempo de resposta com vários parâmetros do sistema, mostrando, inclusive, como os parâmetros do sistema poderiam ser escolhidos para otimização de seu desempenho em termos de tempo médio de resposta.

Observe que este tipo de Plataforma de EaD permite que a largura de banda C da rede seja aumentada. Este parâmetro, por exemplo, pode ser usado para maximizar o desempenho da Plataforma de EaD. Ou seja, a partir de nosso modelo, podemos prever o comportamento da Plataforma quando acessada por certa quantidade de usuários. Portanto, a abordagem proposta pode ser vista como uma poderosa ferramenta para projeto e implantação de Plataformas de EaD do tipo Cliente-Servidor em todo o mundo.

5. Referências

- Alves, L. A., Vieira, F. H. T. e Araújo, S. G. An Analytical Performance Model for Client-Server based e-Learning Platforms. ICELW, 10 a 12 de Junho, Nova Iorque, EUA, 2009.
- Drakopoulos, E. and M. J. Merges, "Performance Analysis of Client-Server Storage Systems", IEEE Transactions on Computers, Vol. 41, No. 11, November 1992.
- Gross, D., Shortle, J. F., Thompson, J. M. e Harris, C. M. Fundamentals of Queueing Theory. 4a Edição, John Wiley & Sons, Nova Iorque, 2009.
- King, P. J. B., Computer and Communication Systems Performance Modelling, Prentice Hall International, UK, 2003.
- Kleinrock, L. Queueing Systems, Volume I: Theory. John Wiley & Sons, Nova Iorque, 1976.
- Pujolle, G. e E. Gelenbe. Introduction to Queueing Networks. 2a Edição, John Wiley & Sons, Nova Iorque, 1998.

Acesso Remoto a Redes de Sensores Sem Fio Usando VPN

Matheus Nascimento¹, Bruno Silvestre¹, Silvana Rossetto²

¹Instituto de Informática – UFG

{matheusnascimento,brunoos}@inf.ufg.br

²Departamento de Ciência da Computação – UFRJ

silvana@dcc.ufrj.br

Abstract. *In this work we present a solution based on VPN to connect private WSN (communicating in IPv6), with computers on the Internet (IPv4).*

Resumo. *Neste trabalho apresentamos uma solução baseado em VPN para oferecer conexão a RSSFs privadas (com comunicação IPv6) com computadores ligados na Internet (IPv4).*

1. Introdução

A crescente demanda pela construção dos chamados “espaços inteligentes” tem fomentado a investigação e o uso combinado de tecnologias de sensoriamento e comunicação sem fio acopladas a dispositivos de tamanho reduzido, com capacidade de processamento e armazenamento de dados, e fonte de energia independente [Estrin et al. 2002].

O desenvolvimento de aplicações baseadas nessas tecnologias requerem a construção de redes específicas — Redes de Sensores Sem Fio (RSSFs) — formadas por dezenas, centenas ou até milhares desses pequenos dispositivos [Yick et al. 2008]. A miniaturização e o enredamento dos sensores possibilitam diferentes ângulos de visão sobre um mesmo evento (por exemplo, o deslocamento de um animal ou a variação de temperatura em uma floresta), permitindo observar ambientes físicos ou monitorar equipamentos com elevado grau de precisão.

Com o avanço das tecnologias aplicadas às RSSFs, uma demanda é construir soluções de software que permitam integrar RSSF com outras redes sem fio e com a Internet de forma a possibilitar, por exemplo, o acesso remoto aos dados coletados pelos sensores por diferentes aplicações. Em cenários mais complexos, é possível projetar o desenvolvimento de mecanismos de colaboração para permitir que os dados coletados pelos sensores possam ser associados a outras informações de contexto, permitindo atuar sobre o ambiente em tempo real (por exemplo, dado o nível de umidade do solo e a previsão meteorológica atualizada, controlar dinamicamente a vazão do equipamento de irrigação).

O passo fundamental para alcançar essa integração das RSSFs é conseguir implementar a pilha de protocolos TCP/IP em dispositivos com severas restrições de recursos computacionais. As primeiras experiências de implementação da pilha TCP/IP em dispositivos de baixo consumo serviram para demonstrar a viabilidade desse esforço [Dunkels 2003, Hui and Culler 2008] e motivar a criação do grupo de trabalho 6LoWPAN (designado pela IETF para propor mecanismos e recomendações básicas para permitir a transmissão de pacotes IPv6 em redes de baixo consumo, entre as quais estão

as RSSFs). O protocolo IPv6 (sucessor do IPv4) tem sido adotado desde o início como protocolo de roteamento padrão para a integração das redes de baixo consumo (LoW-PAN) com a Internet, pois oferece um amplo espaço de endereçamento e características adicionais que facilitam a compactação do seu cabeçalho, bem como a adoção de outras técnicas para redução da complexidade da sua operação em dispositivos com capacidades limitadas.

Entre as implementações existentes das recomendações do grupo 6LoW-PAN para RSSFs, uma das mais importantes é a **BLIP** (Berkeley IP Information) [TinyOS-Blip 2012]. Em [Nascimento et al. 2011], os autores exploraram as possibilidades oferecidas pelo BLIP e mostraram como ligar os nós sensores com a Internet, usando o protocolo IPv6. Entretanto, como ainda hoje a maioria dos computadores e sub-redes conectadas à Internet operam com o protocolo IPv4, é preciso construir soluções alternativas para permitir a interação remota entre computadores externos (operando em redes IPv4) e os nós sensores em uma RSSF.

Neste trabalho, apresentamos uma solução para essa questão permitindo interconectar os sensores de uma RSSF privada a computadores externos de forma transparente. Propomos uma solução baseada em VPN (*Virtual Private Network*), usando os protocolos IPsec [Kent and Seo 2005] e GRE [Farinacci et al. 2000] de forma integrada. Nosso objetivo principal é permitir que aplicações remotas possam interagir diretamente com os nós sensores da RSSF, requisitando dados sensoreados de forma segura.

O restante deste artigo está organizado da seguinte forma. Na seção 2 discutimos as tecnologias usadas para a criação de VPNs. Na seção 3 apresentamos a solução proposta de interação remota entre computadores externos e os nós sensores em uma RSSF. Na seção 4 mostramos os experimentos realizados. Finalmente, na seção 5 apresentamos as considerações finais do trabalho.

2. Tecnologias para criação de VPNs

VPN (*Virtual Private Network*) é uma forma de interligar computadores criando uma rede privada através de uma rede pública. No nosso caso, estamos interessados em permitir que um computador externo utilize a Internet (rede pública) para se comunicar com sensores remotos, como se esse computador estivesse dentro da RSSF particular (rede privada).

A solução proposta consiste em criar uma VPN interligando um computador externo — conectado a uma rede IPv4 — com a RSSF, de forma que seja possível a comunicação direta entre aplicações que executam nesse computador com a aplicação que executa nos nós sensores, como se todos estivessem na mesma rede. Para isso é preciso criar um túnel entre o computador externo e um computador que funciona como porta de entrada para a RSSF. O túnel fica responsável por despachar as informações de um lado a outro de forma transparente. A Figura 1 ilustra essa proposta.

O túnel pode ser realizado em várias camadas da pilha de rede, e geralmente o protocolo do túnel está na mesma camada ou abaixo do protocolo que está sendo encapsulado. O que ocorre é que as informações do protocolo encapsulado passam a ser o *payload* do protocolo de transferência do túnel. Podemos citar como exemplo pacotes IPv6 sendo “tunelados” por uma rede IPv4 (mesma camada), ou pacotes IP sobre IPX (camada abaixo). Como a transferência de dados se dará através da Internet, temos que garantir os requisitos mínimos de segurança na comunicação.

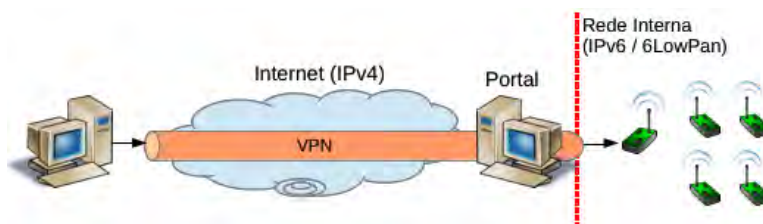


Figura 1. VPN com túnel IPv6/IPv4.

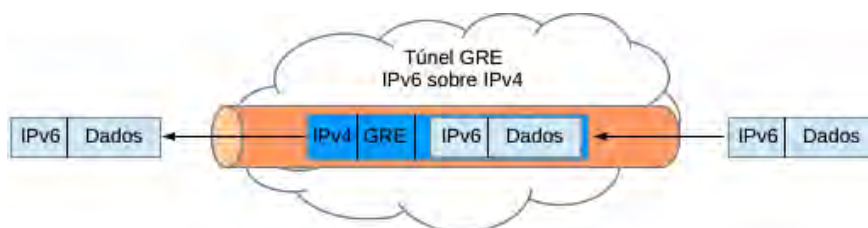


Figura 2. Túnel IPv6/IPv4 utilizando GRE.

2.1. GRE (*Generic Routing Encapsulation*)

GRE [Farinacci et al. 2000] permite o encapsulamento de um protocolo em outro. De forma geral, os dados do protocolo a ser encapsulado torna-se o *payload* do novo pacote. O GRE adiciona um cabeçalho extra contendo informações de controle e finalmente é adicionado ao pacote o cabeçalho do protocolo de transporte dos dados. A Figura 2 mostra o exemplo de um túnel IPv4 transportando pacotes IPv6.

O GRE é utilizado principalmente por sua flexibilidade. O cabeçalho necessita apenas de algumas informações de controle e o identificador do protocolo que está sendo transportado, o que gera um baixo *overhead*. Além disso, o GRE é *stateless*, ou seja, nenhum dos envolvidos necessita manter informações de controle do túnel. No entanto, uma desvantagem é que não há garantia nenhuma de autenticação, integridade ou confiabilidade da informação trafegada, pois o pacote original simplesmente é introduzido como *payload*.

2.2. IPSec

IPsec [Kent and Seo 2005] é uma extensão do protocolo IP que garante segurança nas comunicações em camada de rede. Foi inicialmente desenvolvido para o padrão IPv6 e depois portado para o IPv4.

Ele pode ser utilizado em dois modos: transporte ou túnel — Figura 3 (a) e (b), respectivamente. No modo transporte, apenas os dados do pacote IP original é encriptado ou autenticado. É útil quando apenas se deseja a segurança/autenticação da informação trafegada e é muito usado para comunicação entre duas máquinas. No modo túnel, todo o pacote IP original é encapsulado em um novo pacote IP. Ele permite portanto a criação de uma comunicação segura rede-a-rede, garantindo segurança inclusive dos endereços IPs originais de remetente e destinatário.

O IPsec disponibiliza os protocolos *Authentication Header* (AH) e *Encapsulating Security Payload* (ESP) para garantir autenticação, integridade e confidencialidade dos dados trafegados. O AH é usado para garantir a integridade da comunicação. Ele utiliza



Figura 3. Modo transporte (a) e modo túnel (b) do IPsec.

uma chave secreta para o cálculo de um código de autenticação baseado em *hash*, utilizando diversos campos do pacote IP original. Assim, se o pacote for modificado durante o trânsito, intencionalmente ou não, o receptor ao tentar calcular novamente o código *hash*, muito provavelmente encontrará um código diferente, e então poderá descartar o pacote. O ESP garante confidencialidade dos dados por meio de algoritmos de encriptação, e pode também garantir autenticação por meio de técnicas de *hash*, como no protocolo AH. Diversos algoritmos podem ser utilizados para a criptografia, porém os mais comuns são DES, tripe-DES, AES e Blowfish. Há ainda o algoritmo NULL que não encripta os dados, útil para depuração.

Para o estabelecimento do canal seguro, o IPsec precisa que os envolvidos troquem um conjunto de parâmetros. Assim, o estabelecimento da conexão (chamada de *Security Association* (SA)) foi dividida em duas fases. Na primeira fase, é realizada a identificação dos dois pontos envolvidos e é criado um canal seguro utilizando, por exemplo, uma chave compartilhada entre eles. Esse canal seguro servirá para a fase dois, quando os envolvidos poderão então trocar os parâmetros para configurar o canal definitivo (geralmente os dados necessários para o funcionamento do AH e o ESP).

3. Solução com GRE e IPsec

Nossa solução tem como principais requisitos:

- **Segurança:** como a comunicação se dará pela Internet, é importante que os dados trafegados sejam confidenciais. Além disso, temos que garantir que somente o dono do experimento possa ter acesso, necessitando assim de um mecanismo de autenticação.
- **Transporte por IPv4:** apesar do protocolo IPv6 ser mais apropriado para RSSF, e já contando com implementações no Contiki e TinyOS, muitas instituições não dispõem desse tipo de conexão, recaindo no IPv4. Assim, decidimos focar no transporte de pacotes IPv6 sobre IPv4.

Nenhum dos protocolos descritos acima, atuando isoladamente, consegue preencher esses requisitos. O IPsec possui mecanismos para garantir o item de segurança, entretanto, segundo [Kent and Seo 2005], só é possível realizar associação entre endereços IPv4-IPv4 ou IPv6-IPv6, mas não IPv4-IPv6 — insuficiente para o segundo item.

Por outro lado, com GRE é possível transportar pacotes IPv6 na rede IPv4, mas não há nenhum suporte para segurança. A solução adotada neste caso foi a utilização dos dois protocolos em conjunto. Empregamos o IPsec na base para enviar de maneira segura pacotes IPv4 para o destino e garantir a autenticação na criação da VPN. Com o túnel IPv4 montado, utilizamos então o GRE para encapsular os pacotes IPv6 em IPv4, que serão então enviados via IPsec.



Figura 4. Arquitetura básica IPv4 para montagem de VPN via IPSec.

A seguir apresentaremos como foi realizada a configuração para acesso à rede. Usaremos como teste duas máquinas com Ubuntu 10.04LTS.

3.1. Configuração do IPSec

Na realização de nosso experimento de configuração, criamos um cenário com dois computadores em redes diferentes interligados por um roteador. Como mostrado na Figura 4, temos uma máquina com o endereço 192.168.56.1/24 que deseja acessar a RSSF, e a máquina que tem acesso à RSSF com o endereço 192.168.57.1/24.

Primeiro, precisamos das ferramentas de controle do IPSec. Para isso temos que instalar os pacotes *ipsec-tools* e *racoon* nos dois computadores. Para a configuração da VPN é necessário modificar três arquivos. Apresentaremos a configuração para a máquina externa (192.168.56.1), e a configuração da outra é análoga, trocando apenas os endereços IPs.

O primeiro arquivo a ser configurado é o `/etc/ipsec-tools.conf`, que cria a política de comunicação de entrada e saída de pacotes entre as duas máquinas, dizendo que essa comunicação utilizará o protocolo ESP (para criptografia dos pacotes) em modo transporte:

```
1 spdadd 192.168.56.1 192.168.57.1 any -P out ipsec
2   esp/transport//require;
3 spdadd 192.168.57.1 192.168.56.1 any -P in ipsec
4   esp/transport//require;
```

Além de informar que a política de comunicação ponto-a-ponto entre as máquinas é segura, temos que configurar o sistema com os parâmetros para efetivamente criar esse canal seguro. Adicionamos então no arquivo `/etc/racoon/racoon.conf` a seguinte configuração:

```
1 path pre_shared_key "/etc/racoon/psk.txt";
2 remote 192.168.57.1 {
3     exchange_mode aggressive;
4     proposal {
5         encryption_algorithm aes;
6         hash_algorithm sha1;
7         dh_group modp1024;
8         authentication_method pre_shared_key;
9     }
10 }
11 sainfo address 192.168.56.1 any address 192.168.57.1 any {
12     pfs_group modp768;
13     encryption_algorithm aes;
14     authentication_algorithm hmac_md5;
15     compression_algorithm deflate;
16 }
```

Essa configuração define como serão realizadas as duas fases de estabelecimento de conexão do IPsec. A diretiva *remote* controla a primeira fase, dizendo que a conexão com a máquina conectada à RSSF utiliza criptografia AES, com hash SHA1 e o protocolo Diffie-Hellman (DH) — linhas 5, 6 e 7. O protocolo DH utiliza chave simétrica, e a linha 8 informa que essa chave deve ser buscada no arquivo de chaves compartilhadas, definido na linha 1. Como a mesma configuração será feita nas duas máquinas, ambas terão a mesma chave e o protocolo DH funcionará corretamente. A diretiva *sainfo* possui os parâmetros para a segunda fase do IPsec entre as duas máquinas. A linha 12 configura o PFS¹, que é uma propriedade de troca de chaves simétricas usada na conexão. Temos então a indicação do uso dos algoritmos AES e HMAC-MD5 nas linhas 13 e 14. Finalmente, a indicação de que não há compressão de dados, descrito na linha 15.

Por fim, temos que informar a chave no arquivo `/etc/racoon/psk.txt`. Essa chave pode ser indicada como uma sequência de caracteres ASCII. No nosso caso, geramos um identificador aleatório e o utilizamos nas duas máquinas. Perceba que a chave está associada com o IP da máquina remota:

```
192.168.57.1    512fecfa-064c-43e8-8fda-94d0aba1d4ab
```

Temos que carregar nossa configuração executando os seguintes comandos em um *shell* como *root* — eles irão reiniciar os *daemons* responsáveis pela VPN.

```
# service setkey restart
# service racoon restart
```

3.2. Configuração do GRE

Com a conexão IPsec configurada, iremos criar um túnel IPv6-IPv4 entre as duas máquinas. Novamente, mostraremos a configuração de uma das máquinas, a configuração da outra é análoga. Assumindo o prefixo IPv6 `FEC0::/64` para a RSSF.

Primeiramente, temos que carregar o módulo do *kernel* chamado *ip_gre*, que é responsável pela gerência do túnel. Em seguida, vamos configurar o túnel, o que dá origem a uma nova interface virtual de rede no sistema (*gretun*) e a colocamos ativa. Finalmente, configuramos um endereço IPv6 que será utilizada na comunicação com a RSSF. Os comandos devem ser digitados no *shell* como *root*.

```
# modprobe ip_gre
# ip tunnel add gretun mode gre remote 192.168.57.1
  local 192.168.56.1 ttl 64 dev eth0
# ip link set gretun up
# ip -6 addr add 2000::1/64 dev gretun
```

A Figura 5 mostra a nossa rede virtual IPv6 sobreposta à rede IPv4 (não mostrada). A máquina ligada à RSSF servirá como um roteador, repassando os pacotes da RSSF para a máquina externa. Nesse caso, é interessante configurar na máquina externa (somente nela) a rota para a RSSF via `2000::2`:

```
# ip -6 route add FEC0::/64 via 2000::2
```

¹*Perfect Forward Secrecy*: novas chaves simétricas não serão derivadas da chave da fase 1.

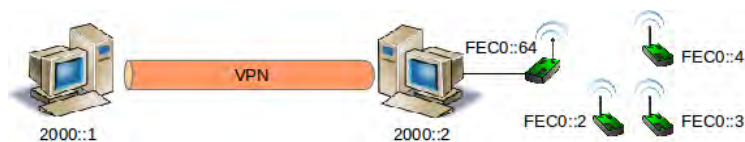


Figura 5. Túnel IPv6/IPv4 com GRE e IPSec para acesso à RSSF.

4. Avaliação

Para avaliar a nossa solução, criamos um cenário com três computadores (como ilustrado na Figura 4): um simulando a rede externa, outro com acesso à RSSF e o terceiro sendo usado como roteador. Realizamos toda a configuração já descrita, obtendo a configuração da Figura 5. Todos os computadores executavam Ubuntu 10.04 TLS.

Utilizamos a versão 1.0 do BLIP (pilha IPv6 para RSSF) que está disponível no TinyOS 2.1.1. O nó sensor com endereço FEC0::64 funciona como nó *sink*, e instalamos nos demais a aplicação de demonstração UDPEcho, disponível no TinyOS. Essa aplicação implementa o serviço Echo, que basicamente envia de volta o pacote UDP recebido na porta 7. Da máquina externa foi possível interagir com os nós sensores, enviando pacotes (contendo strings) com o programa *nc* (*netcat*), e recebendo-os de volta.

Em um exemplo mais complexo, utilizamos a aplicação apresentada em [Nascimento et al. 2011], que permite interagir com os sensores por meio de um navegador web. A aplicação foi desenvolvida para funcionar como um servidor web limitado, sendo executado no nó sensor, e permite efetuar a leitura dos sensores ou mesmo acender ou apagar alguns *leds* através de requisições HTTP. Para essa interação com os nós sensores foi utilizando o navegador Firefox, o qual estava sendo executado na máquina externa.

5. Considerações Finais

A interligação das RSSFs com a Internet começa a virar realidade com a utilização do protocolo IPv6 por parte dos nós sensores. Por outro lado, não temos ainda a adoção desse protocolo em massa na Internet, impossibilitando o acesso direto às RSSFs executando em IPv6. Devemos então utilizar soluções transitórias, por exemplo, o protocolo IPv4 para transportar IPv6. Além disso, nem toda RSSF deve ser acessível publicamente. Assim como algumas redes corporativas, que só liberam o acesso a seus serviços via VPN, podemos ter cenários de RSSF privadas (por exemplo, “casas inteligentes”) onde o acesso através da Internet seja possível, mas controlado.

Apresentamos neste trabalho uma solução baseada em IPSec e GRE para acesso à RSSF considerada privada. A utilização dos dois protocolos foi necessária pois ambos, isoladamente, não possuíam os requisitos para disponibilizar o acesso remoto. Utilizando o IPSec foi possível criar um canal seguro entre a máquina remota e um computador com acesso à RSSF. Esse canal, além de manter a comunicação criptografada, utiliza uma chave para estabelecimento da conexão, o que permite controlar quem está requisitando acesso. Já o GRE foi empregado para efetuar o transporte de pacotes IPv6 sobre IPv4.

A configuração do GRE não apresenta grande dificuldade, no entanto, o protocolo IPSec é bem mais complexo. O IPSec é dividido nos protocolos AH e ESP, que oferecem

várias funcionalidades, sendo necessário diversos parâmetros de configuração (por exemplo, modo de funcionamento, algoritmos de criptografia, compressão e chaves). A curva de aprendizado se mostrou muito alta, ainda que, em nossa opinião, exploramos pouco os protocolos. Além disso, a configuração sempre se baseia em modificações de arquivos de configuração e recarga de *daemons*. A construção de sistemas, seja via web ou desktop, seria mais simples se houvesse uma forma mais simples de interagir com o SO, ou seja, uma API de mais alto nível para a programação.

O ponto positivo dessa solução é a utilização de protocolos padronizados, por outro lado, a configuração do IPSec pode se tornar cada vez mais complexa dependendo do cenário que ele for empregado. Gostaríamos de avaliar, em trabalhos futuros, outras alternativas como o OpenVPN [OpenVPN Inc. 2012] ou o PPTP [Hamzeh et al. 1999].

6. Agradecimentos

Este trabalho foi desenvolvido como parte do projeto CIA2², financiado pela Rede Nacional de Ensino e Pesquisa (RNP).

Referências

- Dunkels, A. (2003). Full TCP/IP for 8-bit architectures. In *Proceedings of the 1st International Conference on Mobile Systems Applications and Services (MobiSys)*, pages 85–98. ACM Press.
- Estrin, D., Culler, D., Pister, K., and Sukhatme, G. (2002). Connecting the physical world with pervasive networks. *Pervasive Computing, IEEE*, 1(1):59–69. <http://doi.ieeecomputersociety.org/10.1109/MPRV.2002.993145>.
- Farinacci, D., Li, T., Hanks, S., Meyer, D., and Traina, P. (2000). RFC2784: Generic routing encapsulation (GRE). <http://tools.ietf.org/html/rfc2784>.
- Hamzeh, K., Pall, G., Verthein, W., Taarud, J., Little, W., and Zorn, G. (1999). RFC-2637: Point-to-point tunneling protocol (PPTP). <http://www.ietf.org/rfc/rfc2637.txt>.
- Hui, J. and Culler, D. (2008). IP is dead, long live IP for wireless sensor networks. In *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems (SenSys)*, pages 15–28. ACM Press.
- Kent, S. and Seo, K. (2005). RFC4301: Security architecture for the internet protocol. <http://tools.ietf.org/html/rfc4301>.
- Nascimento, M., Silvestre, B., and Zenha, L. (2011). Interligando RSSFs com a Internet utilizando IPv6. In *Encontro Anual de Computação*.
- OpenVPN Inc. (2012). OpenVPN. <http://www.openvpn.net>.
- TinyOS-Blip (2012). Blip — Berkeley IP Information. <http://smote.cs.berkeley.edu:8000/tracenv/wiki/blip>.
- Yick, J., Mukherjee, B., and Ghosal, D. (2008). Wireless sensor network survey. *Computer Networks*, 52:2292–2330.

²Construindo Cidades Inteligentes: da instrumentação dos ambientes ao desenvolvimento de aplicações — <http://www.nr2.ufpr.br/~cia2>

O Uso da Tecnologia de Informação em Sala de Aula: Um Estudo de Caso.

Waleska Miranda Lima¹, José Elenilson Cruz²

¹Acadêmica da especialização em Gestão Pública Municipal– Universidade Estadual do Goiás (UEG) – Formosa – GO – Brazil

²Orientador do Curso de Pós-graduação *lato sensu* em Gestão Pública Municipal– Universidade Estadual do Goiás (UEG) – Formosa– GO – Brazil

waleskamlima@gmail.com, elenilsoncruz@hotmail.com

Abstract. *Whereas the use of information technologies and of utmost importance to educators in the classroom, this article aimed to discuss, from the perspective of the teacher, that way the use of these technologies favor the work carried out in the classroom, and what are the barriers faced by the educator, to make this a reality. Classifying them is an exploratory qualitative, this work makes a case study in the school Education Center 10 of Ceilândia, an institution benefited, in 2012, by program UCA (one computer per student) of the Federal Government. It was found that among the benefits provided by the use of information technology in the classroom are the cognitive development and the formation of knowledge. In addition, students learn how to conduct research, prepare spreadsheets, calculations and drafting texts, for example. However, there are factors that make the best use of technology, such as the availability of machines, technological infrastructure, low internet speed, physical infrastructure, and lack of knowledge on the part of teachers.*

Resumo. *Considerando que o uso das tecnologias de informação é de suma importância aos educadores em sala de aula, este artigo objetivou discutir, na perspectiva do professor, de que forma o uso dessas tecnologias favorecem os trabalhos realizados em sala de aula, e quais são as barreiras enfrentadas pelo educador, para tornar isto uma realidade. Classificando-se exploratório-qualitativo, este trabalho realiza um estudo de caso na escola Centro de Ensino 10 da Ceilândia, uma instituição beneficiada, em 2012, pelo programa UCA (Um computador por aluno) do Governo Federal. Verificou-se que dentre os benefícios proporcionados pelo uso das tecnologias de informação em sala de aula se destacam o desenvolvimento cognitivo e a formação do conhecimento. Ademais, os alunos aprendem a realizar pesquisas, elaborar planilhas, cálculos e redigir textos, por exemplo. Entretanto, há fatores que dificultam o melhor aproveitamento da tecnologia, como a disponibilidade de máquinas, infraestrutura tecnológica, baixa velocidade da internet, infraestrutura física e falta de conhecimento por parte dos professores.*

1. Introdução

A evolução das tecnologias digitais vem favorecendo o crescimento do uso das ferramentas de comunicação e expressão (blogs, programas de conversas instantâneas, grupos de relacionamentos etc.) em sala de aula, e as consolidando como eficazes instrumentos pedagógicos. Essas ferramentas, quando bem utilizadas, oferecem a todos educadores inúmeras possibilidades de enriquecer o processo de ensino e aprendizagem.

Com esse pensamento, e com a popularização da internet, cresce de maneira positiva a aplicação das tecnologias de informação em projetos educacionais. As tecnologias de informação são importantes recursos didáticos a serviço dos professores, de uma forma geral, porque promove maior interatividade com o estudante, dentro e fora de classe, enriquecendo, de maneira especial, o processo de ensino e fortalecendo o contato pedagógico entre aluno e professor.

Nesse sentido, embora sabendo dos benefícios que as tecnologias trazem ao ambiente escolar, e ciente de que o uso desses instrumentos como estratégias didáticas podem representar um grande desafio para o educador, questiona-se se há e quais são as dificuldades enfrentadas pelo educador em relação ao uso da tecnologia de informação no ambiente de sala de aula?

Nessa perspectiva, a justificativa para a realização deste estudo reside na importância de se mostrar os benefícios da tecnologia da informação, para o aprimoramento da didática do educador em sala de aula, mas também identificar se há barreiras nesse processo.

Assim, o objetivo geral deste trabalho é discutir, na perspectiva do professor, de que forma o uso das tecnologias favorece os trabalhos realizados em sala de aula, e quais as barreiras enfrentadas pelo educador, para tornar isto uma realidade. Como objetivos específicos, o estudo visa discutir sobre o uso das tecnologias de informação em sala de aula e estimular o educador quanto ao uso da tecnologia em sala de aula.

Este trabalho está dividido em oito partes, sendo: a introdução, que apresenta o tema, a pergunta de pesquisa, hipóteses e objetivos; a diferença de informática educativa e informática educacional, o referencial teórico, composto pelos aspectos que dizem respeito à informática educacional, e o uso das tecnologias em sala de aula no Brasil; a metodologia, que discorre sobre os métodos de pesquisa aplicados; a discussão e os resultados, e as considerações finais, que tecem as principais conclusões do trabalho articulando-as com os objetivos e o problema de pesquisa.

2. Definição de Informática Educativa e Informática Educacional

Há diferença entre Informática Educativa e Informática Educacional, e em alguns momentos algumas pessoas podem se confundir com as duas expressões. A primeira está vinculada à visão educacional e a segunda a condição pedagógica em que o computador é utilizado.

Na informática educativa o computador é utilizado como um material didático, ou seja, como mecanismo de auxílio mínimo no processo de ensino e aprendizagem, não precisando de planejamento prévio.

Assim, a informática educacional adota-se a definição de que o computador é um instrumento primordial no processo de ensino e aprendizagem, com a necessidade de planejamento prévio e a integração entre as disciplinas, para que se possa utilizar de maneira eficiente todas as possibilidades oferecidas, entre outros. O educador pode adaptar os conteúdos vistos de forma teórica em sala de aula para uma prática integrada e ativa no laboratório de informática, por exemplo.

3. Histórico da Informática Educacional

No Brasil, a informática educacional teve seus primeiros avanços no começo de 1971. Em 1981 e 1982, ocorreram em solo brasileiro dois seminários internacionais que tinham como tema principal a informática na educação. Os resultados alcançados por esses seminários acarretaram no surgimento da ideia de implantação de projetos-pilotos nas universidades brasileiras.

Após estes avanços, e apoiado nos resultados obtidos com todas as ações até então, o MEC cria o Programa Nacional de Informática na Educação (PRONINFE) que visava promover o desenvolvimento da informática educativa e seu uso em todo o sistema público de ensino. O PRONINFE ajudou a fomentar o surgimento de infraestrutura de suporte nas escolas, apoiando a criação de centros, subcentros e laboratórios nas escolas; capacitação contínua e permanente dos professores.

Todos os objetivos e metas do PRONINFE foram formulados em sintonia com a política nacional de ciência e tecnologia da época. Nesse sentido, o PRONINFE apresentou os seguintes resultados no período de 1980 à 1995: 44 centros de informática na educação implantados, a maioria interligada na Internet; 400 subcentros implantados, a maioria por iniciativas de governos estaduais e municipais, a partir do modelo de planejamento concebido, inicialmente, pelo Projeto EDUCOM/UFRGS (destes, 87 estão no Estado do Rio Grande do Sul); 400 laboratórios de informática educativa em escolas públicas, financiados por governos estaduais e municipais; mais de 10.000 profissionais preparados para trabalhar em informática educativa no país, incluindo um número razoável de pesquisadores com cursos de mestrado e doutorado.

Foi a partir deste período, que a informática educativa se firmou nas escolas públicas brasileiras. Desde os primeiros projetos que visavam a sua implantação, a cultura de informatização foi se consolidando nas escolas. Apesar das dificuldades orçamentárias, o PRONINFE gerou uma cultura nacional de informática educacional centrada na realidade das escolas públicas, constituindo assim o principal referencial das ações planejadas pelo MEC, correspondendo praticamente a uma fase piloto que durou mais de uma década. Por necessidades burocráticas o programa PRONINFE foi substituído pelo PROINFO em abril de 1997, que na prática era o mesmo PRONINFE, entretanto, com um número muito maior de escolas atendidas. O grande avanço foi a criação de uma estrutura que podia atender os professores que receberam os laboratórios de informática, com capacitação pedagógica e técnica por meio de professores multiplicadores em um Núcleo de Tecnologia Educacional (NTE). Com este suporte técnico e pedagógico para as escolas e para os Núcleos de Tecnologia Educacional (NTE's), uma equipe de suporte pedagógico e técnico em cada NTE era responsável por um número de escolas com laboratórios de informática dentro da sua abrangência geográfica.

O PROINFO teve maior incentivo financeiro e está sendo, até o momento, o mais abrangente programa de informática educativa no território nacional entre todos os projetos, por meio de seus Núcleos de Tecnologia Educacional (NTE). São vários os NTE por Estado, que pesquisam, criam projetos educacionais envolvendo as novas tecnologias da informática e capacitam professores utilizando como suporte os computadores distribuídos em escolas públicas.

Assim, infere-se do exposto que as ações voltadas para a inserção das tecnologias como recursos didáticos não são recentes e vêm se aperfeiçoando na medida em que as estruturas e bases educacionais envolvidas com estes processos se solidificam, a começar pelo entendimento da informática educativa e/ou informática educacional.

4. Uso das Tecnologias em Sala de Aula no Brasil

O Comitê Gestor da Internet no Brasil (CGI.br), setor do Núcleo de Informação e Coordenação do Ponto BR, responsável por coordenar e integrar as iniciativas e serviços da internet no país, divulgou neste ano de 2012, a segunda pesquisa realizada e que teve por objetivo analisar o uso do computador e da internet pelas escolas brasileiras. Foram analisados 650 estabelecimentos educacionais, sendo 497 escolas públicas e 153 particulares.

Os resultados trazidos pela pesquisa evidenciaram que na rede pública de ensino o uso do computador e da internet ainda é considerado como um desafio, mesmo diante dos investimentos que o Governo tem realizado na área de tecnologia da informação. A pesquisa evidenciou que do total de 6.364 alunos entrevistados, 95% já fizeram uso do computador e 91% já usaram a internet, independentemente do local (NIC. BR, 2011).

Para complementar, a pesquisa demonstrou o local mais frequente em que os alunos acessam a internet, conforme demonstra a tabela 1 adiante.

Tabela 1: Local em que os alunos mais acessam a internet

Percentual (%)		Em casa	Escola Pública	Escola Privada	Na casa de outra pessoa ³	Local público de acesso gratuito ⁴	Local de acesso pago ⁵	Em outro local
Total		63	7	2	10	16	1	1
SEXO	Feminino	62	7	2	11	16	1	1
	Masculino	63	7	2	10	16	1	1
REGIÃO	Norte / Centro-Oeste	56	10	2	12	18	1	1
	Nordeste	51	6	2	14	26	1	0
	Sudeste	71	6	2	9	10	1	1
	Sul	69	11	1	9	8	1	1
DEPENDÊNCIA	Pública Municipal	47	13	2	14	22	2	0
	Pública	60	5	2	13	18	0	2

ADMI NISTR ATIVA	Estadual							
	Total – Pública	53	10	2	13	20	1	1
	Particular	87	2	1	5	5	0	0
SÉRIE	4ª série / 5º ano do Ensino Fundamental	52	14	1	15	16	2	0
	8ª Série / 9º ano do Ensino Fundamental	66	5	1	9	18	1	0
	2º ano do Ensino Médio	70	3	2	9	14	0	2

Conforme demonstrado na tabela 1, do total de alunos entrevistados apenas 7% destes afirmaram que acessam a internet na escola, contra 63% que declararam realizar este acesso em casa. O dado é ainda mais espantoso quando visualizado o percentual de alunos que acessam a internet na rede privada de ensino, representando apenas 2% destes. A informação pode levar a crer que a rede pública está mais bem preparada para inserir e oferecer as tecnologias de informação do que a rede particular.

Apesar de a rede pública possuir maior disponibilidade de tecnologias de informação, em comparação à rede privada, isso, necessariamente, não significa que as escolas públicas detenham maior nível de qualidade. Falta, por parte dos educadores, conhecimento em relação a uso da tecnologia, há dificuldades relacionadas à disponibilidade de máquinas para todos os alunos e também restrições quanto à velocidade da internet.

Desse modo, aspectos relacionados à falta de conhecimento por parte dos professores e ainda limitações quanto à disponibilidade de máquinas e baixa velocidade da internet são fatores dificultadores que devem ser considerados para a melhoria do processo de ensino/aprendizagem de informática educacional em sala de aula.

Diante dos dados e argumentos apresentados, pode-se afirmar que embora a informação e a comunicação sejam atualmente imprescindíveis a alunos e professores, na rede pública de ensino o uso do computador e da internet ainda é considerado como um desafio a ser superado, principalmente quando se analisa as dificuldades de interação entre os usuários e softwares.

5. Metodologia da Pesquisa

Este estudo efetuou a pesquisa exploratória, para mostrar os fatos e fenômenos da tecnologia em sala de aula e caracterizar o objeto de estudo e definir a amostra, bem como mencionar os instrumentos de coleta de dados utilizados.

Os dados foram coletados por meio de pesquisa bibliográfica em artigos, monografias e teses para a revisão de importantes aspectos teóricos e coleta de dados secundários sobre o tema.

6. Objeto de Estudo e Instrumentos de Coleta de Dados

O estudo foi realizado no Centro de Ensino 10 da Ceilândia, uma escola pública de Ensino Fundamental, situada à EQNN 23/25, em Ceilândia Norte, Distrito Federal. Esta escola possui 514 alunos e 42 professores. Destes docentes, apenas 4 atuam diretamente

com a tecnologia em sala de aula. Sua escolha se deu em razão de possuir laboratório e ter sido beneficiada, neste ano, pelo programa UCA (Um computador por aluno) do Governo Federal.

Como a investigação deste trabalho é na perspectiva do professor, as perguntas formuladas sobre o uso de computadores e internet em sala de aula objetivaram investigar os problemas enfrentados pelo professor e a sua percepção quanto aos benefícios trazidos pelo uso de tecnologias da informação como procedimento pedagógico. Para a efetivação do estudo de caso, os dados primários foram coletados por meio de roteiro da entrevista aberta com dois professores do Centro de Ensino 10 da Ceilândia, o qual respondeu livremente. Cabe ressaltar que a escola possui apenas quatro professores que atuam com o uso da tecnologia em sala de aula, e um destes encontra-se atualmente de férias, motivo pelo qual a entrevista não foi realizada também com este profissional.

7. Resultados e Discussão

A escola possui em média 470 computadores, dos quais apenas 235 estão em uso. Os computadores têm baixo uso pelos alunos, apenas uma vez por semana. Um dos grandes problemas enfrentado pela escola atualmente é dispor de um local para guardar os computadores, uma vez que se trata de laptops fornecidos pelo Governo Federal através do programa “Um Computador por Aluno”. Embora os computadores estejam em bom estado de conservação, a ausência de armários para guardá-los prejudica o uso do equipamento diariamente, uma vez que o processo de retirada e guarda de tantos equipamentos em local adverso tem comprometido o horário das aulas.

Como se observa, mesmo com os investimentos em tecnologias realizadas pelo Governo, simples problemas de ordem estrutural ainda precisam ser superados. Conforme argumentado anteriormente, a partir dos dados da pesquisa realizada pelo Comitê Gestor da Internet no Brasil, o uso do computador e da internet ainda é considerado como um desafio nas escolas brasileiras.

Além deste problema, a internet na escola é muito lenta e não suporta o elevado número de acessos que são realizados simultaneamente. Diante dessa dificuldade, os alunos são orientados a utilizar o computador em forma de rodízio, para permitir que todos usem o equipamento.

Apesar de tantos problemas enfrentados, os benefícios proporcionados pelo uso de computadores e internet em sala de aula são muito positivos e perceptíveis, uma vez que os alunos demonstram interesse e aprendem rapidamente. Dentre os principais benefícios proporcionados pelo uso de tecnologias da informação na escola, se destacam o desenvolvimento cognitivo e a formação do conhecimento. Os alunos aprendem a realizar pesquisas, elaborar planilhas, cálculos e redigir textos, por exemplo.

Um fator positivo ainda evidenciado foi o avanço dos alunos no uso e manuseio do computador. Os professores admitem que atualmente as crianças e jovens são muito mais avançados no uso da tecnologia do que muitos professores. Nesse sentido, vale a pena lembrar que na pesquisa realizada pelo CGI.br, 95% dos alunos já fizeram uso do computador e que 91% já utilizaram a internet, e que o principal local de uso desta ferramenta é a residência.

Por fim, os professores declararam que atualmente possuem domínios dos principais programas utilizados como Word, Excel e PowerPoint, embora no início não soubesse manusear corretamente o computador. Reconhecem que a capacitação continuada é fator determinante para o sucesso das aulas e principalmente para o favorecimento do processo de ensino e aprendizagem. Os professores concluíram sua declaração refletindo que conhece alguns professores mal sabem ligar um computador. Assim, conforme mencionado anteriormente, a falta de conhecimento por parte dos professores e as limitações tecnológicas são de fato aspectos que dificultam o processo de ensino/aprendizagem de informática educacional em sala de aula.

8. Considerações Finais

Diante do que foi discutido, concluiu-se que o uso das tecnologias da informação, ou seja, computador e internet em sala de aula se tornou determinante para o futuro promissor de crianças e jovens em idade escolar. Na escola pesquisada, sob o ponto de vista dos professores, o uso das tecnologias contribui para a formação dos alunos, principalmente no que se refere a formação do conhecimento, desenvolvimento e aperfeiçoamento cognitivo, além de favorecer a organização de dados, a construção de tabelas, esquemas e desenhos, a resolução de cálculos e o processamento e tratamento de textos.

No Brasil, de uma forma geral, o uso dessas tecnologias não tem sido realizado de maneira eficiente nas escolas, uma vez que os alunos declararam, em pesquisa realizada pelo Núcleo de Informação e Coordenação do Ponto BR, que fazem uso do computador com maior frequência em casa e que aprendem a usá-lo também neste ambiente.

Assim, respondendo à pergunta de pesquisa, conclui-se que o educador enfrenta dificuldades em relação ao uso da tecnologia de informação no ambiente de sala de aula. Dentre as quais estão a indisponibilidade de máquinas para todos os alunos, precária infraestrutura física, com a falta de local adequado para a guarda dos laptops, baixa velocidade da internet e pouco conhecimento (de alguns professores) quanto ao uso das tecnologias. Estas dificuldades são fatores que dificultam a interatividade entre alunos e professores, entre professores e máquinas/software, e reduzem a eficiência do processo de ensino-aprendizagem.

Diante do exposto, além das providências quanto à solução dos problemas relacionados à infraestrutura física e tecnológica, é preciso que a escola, governo e os próprios professores busquem a atualização e o aperfeiçoamento contínuo do corpo docente em relação às tecnologias de informação. Apenas esse conjunto de ações podem maximizar os benefícios da tecnologia para o aprendizado e desenvolvimento dos alunos.

9. Referências

ANDRADE, P. F.; Programa Nacional de Informática Educativa. A utilização da Informática na escola pública brasileira. (1970-2004); MEC: Secretaria de Educação a Distância, 2006.

- BORGES, Priscilla. Escolas privadas são menos equipadas que públicas. Disponível em: <<http://ultimosegundo.ig.com.br/educacao/2012-05-20/escolas-privadas-sao-menos-equipadas-que-publicas.html>>. Acesso em: 23 nov. 2012
- CGI - Comitê Gestor da Internet no Brasil. Pesquisa sobre o uso das tecnologias de informação e comunicação no Brasil: TIC Educação 2011 = Survey on the use of information and communication Technologies in Brazil: ICT Education 2011.
- MORAES, Maria Cândida. Informática Educativa no Brasil: uma história vivida, algumas lições aprendidas. In: Revista Brasileira de Informática na Educação. N. 01, set 1997. SÃO PAULO, Secretaria de Estado da Educação.
- MOTA, Emilia Ignez. A história da informática e a sua importância para a unidade escolar. Monografia. Pós-graduação lato sensu em Pedagogia Empresarial. Universidade Cândido Mendes. Rio de Janeiro, 2004.
- NIC.BR - Núcleo de Informação e Coordenação. CGI.br divulga segunda edição da pesquisa TIC Educação. São Paulo, 19 de junho de 2012.
- PROINFO. História da Informática Educativa no Brasil. Disponível em: <<http://edutec.net/Textos/Alia/PROINFO/edprhist.htm>>. Acesso em: 05 nov. 2012.
- TEPETINO, Simone Abichara Santos. A autoformação do professor para uso de tecnologias digitais na educação. Dissertação de Mestrado. Pontifícia Universidade Católica de Minas Gerais. Belo Horizonte, 2004.
- TUFANO, Thiago. Apesar de dificuldades, 93% das escolas públicas têm internet. Disponível em: <http://noticias.terra.com.br/educacao/noticias/0,OI5845803-EI8266,00-Apesar+de+dificuldades+das+escolas+publicas+tem+internet.html>. Acesso em: 22 nov. 2012.

APÊNDICE

Roteiro da Entrevista:

1. A escola possui em média quantos computadores?
2. Todos estão em uso?
3. Com que frequência os alunos se utilizam dos computadores na escola?
4. O professor enfrenta algum problema com o uso de computadores e internet em sala de aula?
5. De que forma os problemas enfrentados são solucionados?
6. Na visão do professor, quais seriam os principais benefícios proporcionados aos alunos pelo uso de tecnologias da informação nas escolas?
7. O professor apresenta domínio em informática? Quais os desafios que enfrentou para se enquadrar nessa nova condição pedagógica?

Monitoramento de Paisagens Urbanas com Redes de Sensores para Auxílio à Decisão em Gestão Participativa

Guilherme P. Gaebler¹, Vinicius S. Patto¹, Marcelo R. Quinta¹

¹Instituto de Informática – Universidade Federal de Goiás (UFG) – Bloco IMF-1
Campus Samambaia – CEP. 74690-815 – Goiânia – GO – Brasil

{guilhermegaebler, viniciussebba, marceloquinta}@inf.ufg.br

Abstract. *The ADGEPA project provide a sensor network to collect environmental data and a responsive web application in which one can make complaints about socioenvironmental concerns of a scenario. From these data, we look for new information using Knowledge Discovery in Database techniques. Our first data polls will be simulated to validate two project modules. In this paper, we present the ADGEPA project, its simulation model, and our Arduino prototype proposal to collect environmental data.*

Resumo. *O projeto ADGEPA prevê uma rede de sensores para captar dados ambientais e uma aplicação web responsiva para usuários fazerem denúncias acerca de aspectos socioambientais em um cenário. Dos dados, pesquisamos por novas informações com uso de técnicas de Descoberta de Conhecimento em Base de Dados. Os nossos primeiros dados serão simuladas para validar dois módulos do projeto. Neste trabalho, nós apresentamos o projeto ADGEPA, o seu modelo de simulação e a nossa proposta de um protótipo para captação de dados ambientais com uso da plataforma Arduino.*

1. Introdução

Muitos dos grandes centros urbanos do mundo apresentam problemas em comum relacionados à mobilidade urbana, poluição, violência, acidentes, saúde pública, falta de infraestrutura adequada para eventos climáticos recorrentes, dentre outros. Com tendências¹ de crescimento constantes, os centros urbanos têm muita carência de formas mais eficientes de gestão do seu espaço físico e de seus recursos socioambientais.

Alguns desses problemas são abordados em pesquisas na esfera da computação. Na equipe de Vikatos [Vikatos et al. 2011] os pesquisadores projetaram um sistema de monitoramento de dados ambientais com o uso de plataformas de sensores. Eles contaram com a participação de voluntários, fizeram testes em ambiente urbano de pequena escala e mediram valores de concentração dos gases metano e monóxido de carbono.

Um exemplo de uso de plataformas de sensores para captar dados ambientais e auxiliar na resolução de problemas sociais é o trabalho de [Rosi et al. 2010], no qual um protótipo para monitoramento de áreas com risco de desmoronamento foi desenvolvido. Consideraram a leitura de dados de movimentos de encostas, temperatura, pressão e umidade.

¹Pode-se constatar essas tendências pelos Censos do IBGE, disponíveis em <http://www.ibge.gov.br/home/>

No trabalho de [Angove et al. 2010], foi desenvolvido um protótipo para a leitura da concentração de monóxido de carbono e compostos orgânicos voláteis (COVs) em ambientes residenciais e comerciais em uma tentativa de associação de problemas de saúde com a exposição a gases. Ward e Wall, por suas vezes, apresentam um sistema de gerenciamento de ambientes internos que, por meio de leituras de sensores e participação dos usuários, auxilia a manutenção de prédios mais sustentáveis, confortáveis e eficientes [Wall and Ward 2009].

Nos trabalhos supracitados, os autores apresentaram propostas de sistemas baseados em sensores para captação, armazenamento e análise de dados ambientais. Eles atacaram problemas bem delimitados e com dados bem definidos. Embora essas contribuições foram muito importantes, acreditamos que seja possível aumentar consideravelmente a amplitude do uso de sensores para um sistema inteligente de gestão de ambientes urbanos de grandes proporções que lidam com diversas variáveis, sejam elas classificáveis pelo senso comum do homem (e.g. existe despejo de esgoto em corpo de água), ou aquelas que exigem equipamentos específicos (e.g. temperatura). Em nosso projeto, identificamos possíveis padrões e relações entre essas variáveis de cunho socioambiental e pesquisamos formas de auxiliar planejamentos e gestões de espaços urbanos.

Apresentamos na Seção 2 uma visão geral sobre o projeto, as motivações, objetivos e metodologias utilizadas. Em seguida, abordamos a simulação no projeto Assistente Digital para Gestão Pública Participativa (doravante, ADGEPA). Posteriormente, na Seção 4, descrevemos em detalhes o nosso modelo de captação de dados ambientais com uso de sensores. Na Seção 5, apresentamos os trabalhos atuais no projeto e as atividades futuras. Por último, apresentamos nossas considerações finais na Seção 6.

2. O Assistente Digital para Gestão Pública Participativa

Observamos que existe um campo pouco explorado no que diz respeito à coleta e análise de dados referentes a aspectos socioambientais em cenários urbanos de médio e grande porte. Sistemas capazes de captar e processar esses tipos de dados e identificar relações e padrões relevantes (a partir desses dados) podem trazer melhores soluções de gestão de recursos socioambientais.

2.1. Motivação

Uma das principais motivações do projeto é a investigação de gestão participativa pela consideração de problemas socioambientais denunciados pela população. A identificação de associações e correlações entre esses problemas pode ser um fator fundamental na escolha de estratégias eficientes para resolução dos principais anseios da população. Um aspecto positivo no processo é o viés participativo e a possibilidade de ser mais eficiente que práticas convencionais de gestão pública (alguns trabalhos já mostraram isso [Adamatti 2007], [Sebba Patto 2010]). Porém, nem todos os fatores socioambientais podem ser percebidos ou são percebidos e classificados de forma imprecisa ou ambígua. Para alguns desses fatores, como luminosidade e umidade relativa do ar, é possível fazer a aferição de valores de forma precisa pelo uso de sensores.

2.2. Objetivos

O projeto ADGEPA propõe uma forma de fazer leituras das percepções de cidadãos sobre determinados problemas socioambientais. A partir dessas percepções, espera-se

auxiliar o poder público na escolha de estratégias que possam minimizar ou eliminar os problemas socioambientais mais relevantes à determinada população.

Com o intuito de coletar dados, pretendemos implementar (i) uma aplicação web responsiva para que usuários façam denúncias sobre problemas socioambientais em cartas geográficas e (ii) um conjunto de miniestações para coleta e envio de dados ambientais. Pelo uso de técnicas de KDD (*Knowledge Discovery in Database*, em português: Descoberta de Conhecimento em Base de Dados) esperamos descobrir informações novas a partir de padrões e associações entre os dados [Fayyad et al. 1996]. A partir das novas informações, pretendemos dar suporte a decisão para escolha de problemas socioambientais a serem atacados.

2.3. Metodologia

Pretendemos fazer coleta de dados socioambientais a partir de duas fontes coletoras: um aplicativo web responsivo e miniestações para coleta de dados ambientais. O aplicativo considera 29 indicadores ambientais e 30 sociais. A plataforma de sensores baseada em Arduino² considera 4 variáveis ambientais (detalhes na Seção 4).

Após a coleta de dados, buscamos padrões e associações entre dados para que possamos compreender melhor possíveis relações entre dados da nossa pesquisa. Em nosso projeto, dispomos do apoio de dois especialistas do domínio que nos ajudam na compreensão dos dados, tanto com relação à granularidade quanto à questão da representatividade de dados. Esses fatores são considerados importantes para a realização de KDD [Silva 2004].

Para tratar as novas informações (descobertas com KDD), modelamos um agente computacional (a ser implementado) que faz uso de técnicas de análise multicritério. Após a escolha de um método de análise multicritério (ainda em estudo), pretendemos apresentar estratégias mais eficientes para resolver os problemas de maior anseio por parte da população (que faz denúncias no aplicativo web responsivo). O gestor artificial terá três perfis e, de acordo com o perfil escolhido, apresentará sugestão para atacar problemas mais relevantes para a população de uma determinada comunidade.

3. Simulação do ADGEPA

3.1. Justificativa

A recolta de dados por parte da aplicação web responsiva e dos protótipos Arduino com sensores ambientais toma muito tempo para que se tenha quantidade de dados satisfatória para buscar novos conhecimentos com KDD. Além disso, seria preciso quantidade elevada de sensores para que eles fossem espalhados em uma cidade. Face a esse grande obstáculo, decidimos simular os dados com uso de sistemas multiagentes. A justificativa para a escolha desta técnica se deve à facilidade de se modelar sistemas orientados a agentes; à possibilidade de se desenvolver modelos que simulem dados coerentes com a realidade; e, além disso, agentes inteligentes possuem características desejáveis para nossa aplicação, e.g.: proatividade, reatividade, cooperação, autonomia e capacidade de operar em ambientes dinâmicos [Wooldridge and Jennings 1995].

²Arduino é uma plataforma open-source para prototipação de dispositivos eletrônicos baseada em hardware e software flexíveis e fáceis de usar. Disponível em: <http://www.arduino.cc/>

3.2. Os agentes

O projeto ADGEPA prevê desenvolvimento de quatro tipos de agentes computacionais.

O primeiro agente é denominado de Agente Ambiente porque ele tem o objetivo de representar ambientes virtuais para Agentes Usuários. A simulação de dados acerca de temperatura, umidade relativa do ar, luminosidade e quantidade de monóxido de carbono deverá ser feita pelo Agente Ambiente.

Os Agentes Usuários são responsáveis pela simulação de atividades humanas em um cenário. Eles farão denúncias de acordo com a percepção do ambiente.

Agentes Mineradores são responsáveis pelo agrupamento temporal e espacial de dados. A partir desses agrupamentos, eles farão busca por associações entre aspectos socioambientais e entre esses aspectos e tempo.

O último agente é denominado Agente Gestor e tem objetivo de auxiliar gestores públicos a encontrar estratégias para reduzir ou eliminar problemas socioambientais mais denunciados pelos usuários do sistema.

3.3. A simulação

Mesmo que o tempo para coletar dados socioambientais via aplicativo web responsivo e via sensores é impraticável, decidimos seguir com o desenvolvimento experimental dos dois artefatos. Porém os dados serão simulados a princípio. Pela simulação multiagentes, acreditamos simular dados coerentes com a realidade, por exemplo, simular poluição de água em um rio ou em um lago (muitos trabalhos de simulação multiagentes apresentaram resultados muito satisfatórios, inclusive para simular humanos, e.g. [Ishida et al. 2007]).

A simulação é um grande desafio, haja vista que é preciso (a) representar um ambiente com dezenas de elementos cartográficos; (b) simular centenas ou milhares de usuários do sistema; (c) garantir que as ações simuladas (pelos Agentes Usuários) sejam coerentes; e (d) garantir que Agentes Usuários se comuniquem como Agente Ambiente de forma que sejam capazes de perceber um trecho de uma carta urbana.

Alguns elementos de cidades são: escolas, igrejas, bares, restaurantes, cinemas, teatros, feiras, supermercados e locais de trabalho. Os Agentes Usuários percebem o mapa pela comunicação com o Agente Ambiente.

Cada simulação representa um intervalo temporal que pode durar meses ou anos; possui um número de Agentes Usuários; e uma representação de um cenário urbano, feita pelo Agente Ambiente.

Findada a simulação da coleta de dados, os Agentes Mineradores são invocados para fazer agrupamentos por regiões (uma região da cidade pode ser dividida em 1000 microrregiões, por exemplo) e momentos (cada uma das microrregiões pode ser agrupada por dias, semanas ou meses). Em seguida, são feitas buscas por associações.

Ao encontrar as associações, entra em cena o Agente Gestor. Este agente faz a seleção das variáveis (denúncias ou leituras de aspectos socioambientais mais relevantes) e procura, de acordo com o seu perfil, quais variáveis devem ser atacadas para minimizar o máximo possível de problemas mais relevantes, segundo as denúncias.

Atualmente, usamos o framework JADE³ por se tratar de uma ferramenta bem aceita por comunidades científicas sobre simulação multiagentes. Além disso, JADE segue os padrões de desenvolvimentos estabelecidos pela FIPA⁴ [Bellifemine et al. 2007].

Na Figura 1, apresentamos graficamente a arquitetura de simulação para o ADGEPA.

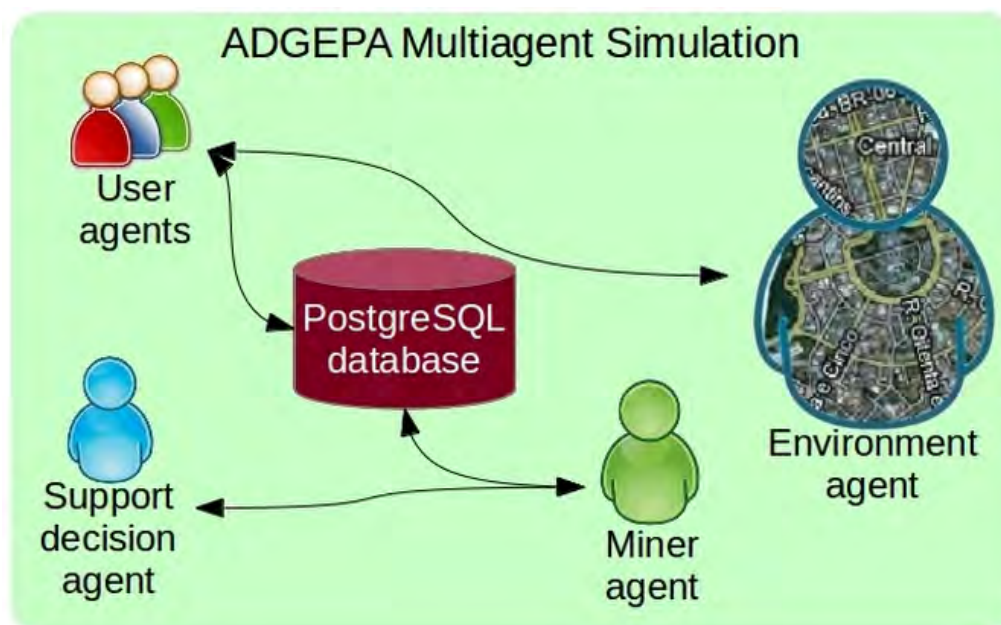


Figure 1. Arquitetura geral da simulação

4. Modelo de captação de dados ambientais

Para a captação de dados ambientais que não são percebidos ou são percebidos e classificados de forma imprecisa ou ambígua por seres humanos, desenvolvemos um mecanismo simples e de baixo custo pelo uso da plataforma de prototipagem eletrônica Arduino. Em nosso protótipo preliminar usamos uma placa Arduino Uno com processador ATmega328P, sensores de umidade, de temperatura, de luminosidade e de concentração de Monóxido de Carbono (CO) no ar, com uma placa Ethernet para a comunicação com o servidor de dados. Experimentalmente, utilizamos como fonte de energia um cabo USB ligado à porta serial do Arduino e uma placa Ethernet para a comunicação. Na Figura 2 apresentamos um esquema do protótipo com o auxílio de uma placa base de prototipação eletrônica (protoboard).

Implementamos um programa em Arduino (linguagem de programação da plataforma homônima). Ele é responsável pela captura e conversão de sinais elétricos dos sensores em informações ambientais. O objetivo desta parte do software é receber e enviar dados pela internet a um servidor web Java, que recebe os dados e os armazena em um banco de dados.

Na versão atual, nosso protótipo afere as seguintes variáveis: temperatura, umidade relativa do ar, luminosidade e concentração de monóxido de carbono. As leituras

³JADE - Java Agent Development Framework. Disponível em <http://jade.tilab.com/>

⁴FIPA - Foundation for Intelligent Physical Agents. Disponível em <http://www.fipa.org/>

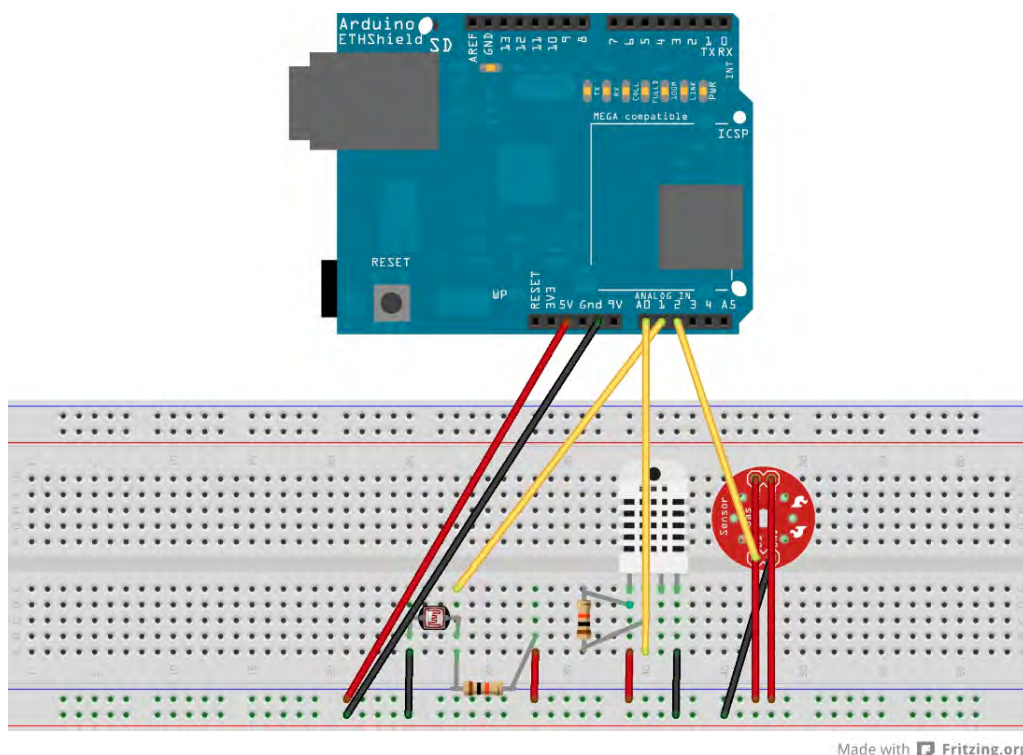


Figure 2. Esquema do protótipo

são feitas em intervalos de tempo pré-definidos. Inicialmente, nosso programa foi ajustado para fazer leituras a cada 10 minutos (esse tempo pode ser facilmente modificado de acordo com a modelagem). A cada leitura, o protótipo verifica se há necessidade de envio de dados ao servidor. Atualmente, descartamos envio de dados quando ocorrem mudanças na ordem de até: dois graus Celsius, cinco por cento da umidade relativa do ar, cinco por cento no valor de luminosidade e cinco por cento no valor da concentração de monóxido de carbono. Os valores desprezados para temperatura e umidade foram escolhidos com base na margem de erro apresentada na especificação do fabricante.⁵ Isto é, se o componente para aferir temperatura atmosférica registra mudança de 1 grau na temperatura (para cima ou para baixo), não é feito novo envio de dados ao servidor. Para os demais componentes, não encontramos informações acerca da margem de erros de aferições nas especificações do fabricante. Assim, propomos desprezar mudanças na ordem de até 5 por cento.

Em nosso modelo, pequenas variações nos valores não causam envio dos dados. Essa decisão proporciona economia de energia (que pode ser um ponto crítico no protótipo de sensores) e a redução de tráfego de dados desnecessários (que pode ser um ponto crítico por se tratar de conexão em redes com capacidade de tráfego limitado).

5. Trabalho atual e prospecto de atividades

O nosso projeto é bastante abrangente e envolve simulação multiagentes; captação de dados socioambientais via aplicativo e via rede de sensores; descoberta de conhecimento; e suporte a decisão. A maior parte dos agentes se encontram em fase de modelagem

⁵Disponível em <http://www.micro4you.com/files/sensor/DHT11.pdf>

(agente usuário, agentes mineradores e agente gestor). O agente ambiente está em fase de codificação. Com relação à plataforma de sensores, estamos em fase de estudo para viabilizar o uso do protótipo; para isso, é preciso que (i) o fornecimento de alimentação elétrica ao protótipo seja mais duradouro e (ii) que a forma de envio de dados não dependa de uma rede Ethernet e possa ser realizada através de rede sem fio.

6. Conclusão

Neste artigo, apresentamos o projeto ADGEPA e demos ênfase em uma das suas formas para captação de dados: uso de dispositivos sensoriais para leitura de dados ambientais. Tendo em vista a necessidade de grande lapso temporal para coleta de dados pela aplicação web responsiva e pelos sensores de captação de variáveis ambientais, o projeto propõe a simulação de dados com o uso de Sistema Multiagentes. O projeto ainda é recente e não possui todos os seus módulos desenvolvidos. Ainda será necessário fazer a integração das partes, testes e avaliações ao final deste processo.

Pelo uso da plataforma Arduino para captação de dados ambientais, percebemos grande flexibilidade em relação a mudanças. Isso nos possibilita, por exemplo, alterar os tipos de sensores (para captar outras variáveis) ou a forma de comunicação com o servidor, que poderia ser substituída por outras tecnologias como Bluetooth e UMTS (3G). As mudanças podem ser feitas facilmente, com baixo custo e não apresentam impactos ou inconsistências no funcionamento do restante do sistema. Dessa forma, o módulo de captação de dados também é flexível e representa boas oportunidades para estudos futuros, além de proporcionar a captura de dados que seres humanos têm dificuldade de percepção precisa, como temperatura, umidade e pressão do ar.

Sabemos que a simulação multiagentes não representará fielmente ambientes e/ou dados gerados por usuários humanos ou captados por sensores. Todavia, acreditamos que, após a integração dos módulos do projeto ADGEPA, os módulos para descoberta de conhecimento e para suporte a decisão independarão do cenário. Isso, ao nosso entender, os tornam muito promissores. Críticas e dados de projetos similares ou relacionados são bem-vindos.

7. Agradecimentos

Agradecemos aos professores Ivanilton José de Oliveira e Patrícia de Araújo Romão; ao aluno de mestrado Leonardo Afonso Amorim; e aos alunos de graduação Alex Pereira Maranhão, Mirelle Cristina Lima Silva e Carlos Henrique Barros Costa por suas valiosas contribuições nas atividades desenvolvidas no projeto. Atualmente, o projeto é financiado pela FAPEG.

References

- Adamatti, D. F. (2007). *Insertion of virtual players in role-playing games group decision support systems: An experiment in the natural resources management domain*. PhD thesis, Escola Politécnica da Universidade de São Paulo.
- Angove, P., O'Flynn, B., Hayes, J., Diamond, D., O'Grady, M. J., and O'Hare, G. M. (2010). Air-quality monitoring for pervasive health. *IEEE pervasive computing*.

- Bellifemine, F., Caire, G., and Greenwood, D. (2007). *Developing Multi-Agent Systems with JADE*. Wiley, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.
- Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*.
- Ishida, T., Nakajima, Y., Murakami, Y., and Nakanishi, H. (2007). Augmented experiment: Participatory design with multiagent simulation. In *International Joint Conference on Artificial Intelligence (IJCAI'07)*.
- Rosi, A., Bicocchi, N., Castelli, G., Corsini, A., Mamei, M., Zambonelli, F., and Berti, M. (2010). Landslide monitoring in the emilia romagna apennines. *IEEE pervasive computing*.
- Sebba Patto, V. (2010). *The use of assistant agents for the analysis and decision making in participatory management*. PhD thesis, UFR d'Informatique, Université Pierre et Marie Curie.
- Silva, M. P. D. S. (2004). Mineração de dados - conceitos, aplicações e experimentos com weka. *Sociedade Brasileira de Computação*.
- Vikatos, P., Theodoridis, E., Mylonas, G., and Tsakalidis, A. (2011). Patrassense- participatory monitoring of environmental conditions in urban areas using sensor networks and smartphones. In *Panhellenic Conference on Informatics*.
- Wall, J. and Ward, J. (2009). Optimal distributed control for sustainable buildings. *IEEE pervasive computing*.
- Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: Theory and practice. *The Knowledge Engineering Review*.

Simulação baseada em autômatos celulares para a evacuação emergencial de pedestres em ambientes com geometria complexa

André P. de Castro¹, Danielli A. Lima¹

¹Instituto Federal do Triângulo Mineiro (IFTM) – Campus Avançado Patrocínio
Av. Lília Terezinha Lassi Capuano nº 255 – 38.740-000 – Patrocínio – MG – Brazil

Abstract. *The security in buildings environments is an important aspect, especially in areas where permanence and movement of individuals are intense. If a host fails to escape in time to a place of risk, by choosing a wrong way or a bad way out, there is always the imminence of injury and even death. Computer-based simulations are important in determining the reliability and safety of a structural model. The use of cellular automata has proven useful in computer simulation, especially by the parallelism inherent in them. The main objective of the research was to propose a simulation model for emergency evacuation based on the theory of cellular automata adapted from a precursor model.*

Resumo. *A segurança em ambientes de edificações é um aspecto importante, principalmente em áreas em que a permanência e a circulação de indivíduos são intensas. Caso uma multidão falhe ao escapar em tempo de um local de risco, pela escolha de um caminho errado ou de uma saída ruim, há sempre a iminência de lesões e até mesmo de morte. Simulações geradas por computador são importantes para a determinação da confiabilidade e segurança de um modelo estrutural. O uso de autômatos celulares tem se mostrado útil na simulação computacional, principalmente pelo paralelismo inerente dos mesmos. O principal objetivo da pesquisa foi a proposição de um modelo de simulação da evacuação emergencial baseada na teoria de autômatos celulares adaptado de um modelo precursor.*

1. Introdução

A natureza humana manifesta a característica de interagir em sociedade, formando aglomerações para propósitos gerais. Rotinas de segurança devem garantir a integridade física de cada indivíduo dentro de estruturas prediais, onde a formação de multidões pode ocorrer devido a fatores desencadeadores de pânico como, por exemplo, incêndios, conflitos e até mesmo atentados terroristas.

Simulações de evacuações emergenciais, baseadas em sistemas computacionais, influenciam na execução de um plano de segurança predial [Burstedde et al. 2001]. As simulações são estruturadas a partir de abstrações de realidades descritas e o estudo delas geram resultados que, quando comparados, delimitam medidas que servem como base para o nivelamento de segurança de uma estrutura [Perez et al. 2002]. A evacuação sob condições físicas e temporais é relevante, visto que, se um indivíduo ou multidão cometem falhas ao tentar escapar em uma situação de pânico, seja por encontrarem obstáculos em sua trajetória, saídas mal elaboradas ou caminhos incorretos; tendem ao envolvimento

de risco a morte. Em caso de empecilhos na saída predial, lesões podem ocorrer, resultantes de conflitos, pisoteamentos e em casos mais graves a morte de pessoas.

Uma ferramenta útil na elaboração de simulações gerenciadas por computador são estabelecidas por modelos matemáticos conhecidos como Autômatos Celulares (ACs). Existem diversas aplicações no uso de autômatos celulares, dentre elas, podemos citar a modelagem de fenômenos naturais, físicos ou biológicos, que seriam muito difíceis de serem modelados pelas equações diferenciais, sendo estas as mais utilizadas nesse tipo de tarefa. Além disso, os ACs possuem alto nível de paralelismo, quando implementados em um hardware paralelo [Oliveira 2003].

2. Objetivo

Este trabalho tem por objetivo a proposição de um modelo de evacuação com alterações lógicas de implementação do cálculo de movimentação de pedestres utilizando-se autômatos celulares bidimensionais. As características principais do modelo bidimensional proposto por [Varas et al. 2007] serão preservadas. A fim de avaliar o modelo sugerido, testes em diferentes situações foram realizados. As simulações fornecem o tempo de evacuação total de uma quantidade de pessoas e a partir de várias simulações é possível obter a média geral do tempo de evacuação.

3. Fundamentação teórica e trabalhos relacionados

3.1. Autômatos celulares

Um AC é composto por um reticulado com uma dimensão d dividido em células ou unidades processadoras, sendo que, cada célula C é representada por um estado. As células modificam seus estados a cada passo de iteração de acordo com uma regra de transição. Podemos aplicar a regra de transição por T passos de tempo para obter a evolução espaço-temporal do reticulado do AC. A regra estabelecida por uma função de transição indica o novo símbolo a ser escrito na célula do reticulado de acordo com seu estado atual e dos estados de suas vizinhas (regra local). Em sua definição mais usual, a atualização dos estados se dá de forma síncrona e utiliza uma regra determinística, isto é, a cada passo de tempo todas as N células do reticulado são atualizadas. No modelo proposto neste trabalho, o sistema de regras é probabilístico, ou seja, é possível alterar o estado de uma célula a partir de uma função de probabilidade, isso significa que as regras de atualização são probabilísticas e, portanto, o estado de uma célula em um determinado instante depende probabilisticamente dos estados de suas células vizinhas no passo de tempo imediatamente anterior [Lima 2010].

A estruturação de um AC unidimensional (1D) é a forma mais estudada. A mudança de estado de uma célula depende de m vizinhas expressas por $m = (2r + 1)$, sendo r o raio do AC [Oliveira 2003]. Para ilustrar um AC unidimensional considere a Figura 1 (a), que aborda uma modelagem conhecida como regra 30 [Wolfram 2002], contendo um reticulado de 6 células sendo que o estado inicial de cada célula é apresentado em $t = 0$. Uma regra binária de raio 1 é aplicada, sendo que a vizinhança de cada célula é formada por três elementos: a própria célula e suas duas vizinhas adjacentes (à esquerda e à direita). Como esse AC é binário (2 estados possíveis), existem 8 diferentes vizinhanças, da 000 a 111. A regra em si é dada pelos 8 bits de saída associados a cada vizinhança possível: 01111000. Na Figura 1 (b) vemos a atualização do reticulado de $N = 6$ por 2 passos de tempo a partir de sua configuração inicial 101110 em $t = 0$. A

cada passo, cada célula do reticulado é atualizada identificando-se sua vizinhança e seu novo estado é dado pelo bit de saída correspondente na regra de transição. Observe como exemplo, a célula de símbolo 1, destacada em $t = 0$, seu próximo estado será 0 em $t = 1$. O reticulado é submetido a condições periódicas de contorno, sendo que a primeira célula é vizinha imediata da última, e vice-versa. Aplicando-se esse procedimento para todas as células do reticulado de forma síncrona, tem-se a nova configuração do reticulado a cada passo de tempo.

Os ACs fundamentados em duas dimensões (2D) são amplamente conhecidos na literatura. ACs bidimensionais também são muito usados para a representação de padrões gráficos, já que sua execução assemelha-se a uma imagem pixelizada. Essa dimensão contempla variadas formas para a consideração da vizinhança de uma célula, sendo que as duas principais estão apresentadas na Figura 1 (c) e (d), tal que o símbolo $\odot$ representa a célula alvo e o símbolo ■ identifica a célula vizinha. Dentre as aplicações de ACs bidimensionais, o modelo mais conhecido é o jogo matemático proposto por [Conway 1970] denominado *Game of Life*.

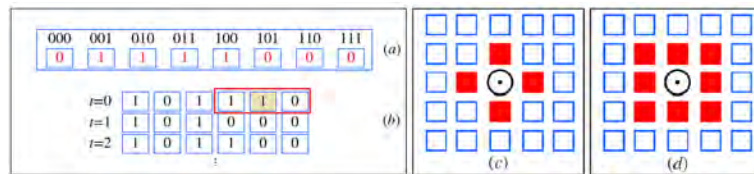


Figura 1. (a) Regra de transição de raio 1. (b) Evolução do AC por $T=2$ passos de tempo. (c) Vizinhança de Von Neumann. (d) Vizinhança de Moore.

3.2. Modelos para evacuação baseados em ACs

Diversos autores apresentaram abordagens de como utilizar ACs para simular ambientes de evacuação de multidões, descrevendo modos de controle do trânsito dos indivíduos e como fatores externos afetam a evacuação. O modelo proposto neste trabalho é fortemente baseado no modelo proposto em [Varas et al. 2007] e [Carneiro 2012]. Os modelos consistem de duas etapas: a primeira etapa, um campo de piso (malha) é demarcado para representar a planta de uma estrutura. A segunda etapa apresenta um conjunto de regras de transição que permitem o deslocamento dos pedestres sobre essa malha. Para isso [Varas et al. 2007] elaborou um modelo de AC bidimensional para simular o processo de evacuação dos pedestres em uma estrutura cujos obstáculos são fixos e estão bem visíveis para os indivíduos que ali se encontram. O campo de piso de uma estrutura é formatado seguindo o conceito de que a partir de qualquer célula é possível seguir o menor caminho até a saída mais próxima. Para simular o comportamento dos pedestres durante a evacuação, em ambos os modelos um fator de pânico foi introduzido, responsável por evitar a movimentação de uma porcentagem de pedestres no ambiente. Segundo [Varas et al. 2007], quando um pedestre deseja se movimentar para uma célula já ocupada, um conflito deve ser resolvido, conforme exemplificado na Figura 2. Por outro lado, em [Carneiro 2012] os pedestres podem seguir rotas alternativas caso uma determinada célula esteja ocupada.

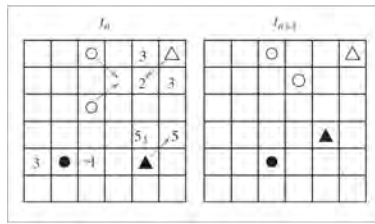


Figura 2. Movimentos possíveis para os pedestres no modelo de [Varas et al. 2007]. Os números correspondem ao valor do campo de piso em uma célula, os arcos representam as intenções de movimento, e os símbolos representam os pedestres em várias situações. Círculo preenchido: o pedestre movimenta-se para uma célula adjacente para minimizar o seu campo de piso. Círculos abertos: dois pedestres tentam se movimentar para a mesma célula; um número randômico decide quem se movimenta. Triângulo preenchido: duas células adjacentes com o mesmo campo de piso, os pedestres decidem randômicamente para onde mover. Triângulo aberto: pedestres em pânico, não se movimentam.

4. Modelo proposto

O modelo proposto é fortemente baseado nos modelos de [Varas et al. 2007] e [Carneiro 2012] e consiste de duas etapas: a primeira etapa é a criação de um campo de piso, onde cada célula desse piso consiste de um valor inteiro e positivo que guia a escolha do pedestre no momento de fuga. A segunda etapa baseia-se utilização de regras que permitem a movimentação dos pedestres de forma segura.

4.1. Campo de piso

Inicialmente, a representação da estrutura predial 3D é transformada para uma estrutura 2D, conforme é observado na Figura 3 (a). Posteriormente, é realizada uma atribuição com valores inteiros em cada célula de piso da estrutura que será simulada. A atribuição dos pesos das células do reticulado segue o algoritmo proposto por [Varas et al. 2007], com a adição do valor da célula adjacente +1.0 para horizontal ou vertical e +1.5 para as células na diagonal. Um exemplo do campo de piso adotado no modelo é apresentado na Figura 3 (c) e (d).

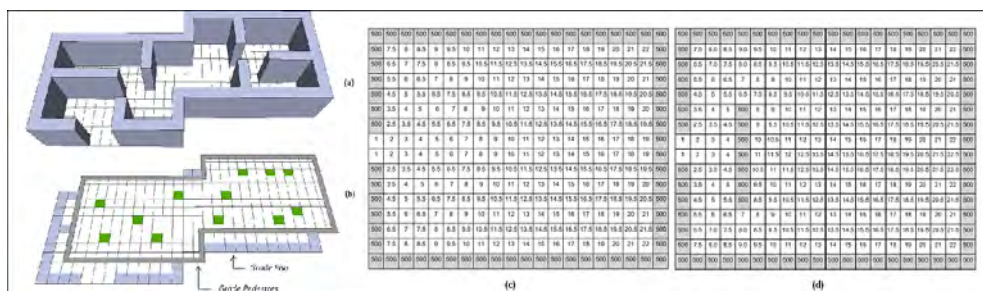


Figura 3. (a) Estrutura predial 3D. (b) Transformação da estrutura 3D para campo de piso 2D e pontos verdes representam os pedestres. (c) Campo de piso com 18×14 células utilizado no modelo [Varas et al. 2007]. A porta de saída está do lado esquerdo (células com valores igual a 1); as células correspondentes às paredes são atribuídas valores de 500. (d) Campo de piso de 18×14 células e um obstáculo na parte central.

4.2. Movimentação de pedestres

Posteriormente uma nova grade com pedestres alocados é considerada no modelo, conforme é mostrado na Figura 3 (b). A alocação inicial desses pedestres não pode sobrepor os obstáculos. Por fim, uma regra que guia a movimentação dos indivíduos na estrutura deve ser elaborada. A regra de movimentação dos pedestres é o principal aspecto que distingue o modelo proposto dos modelos precursores em [Varas et al. 2007] e [Carneiro 2012]. A técnica considerada neste trabalho é descrita a seguir:

A cada passo de tempo os indivíduos no reticulado de pedestres escolhem seu caminho e avançam em direção a saídas. Para a escolha de rota o pedestre sempre tenta ir em direção a porta mais próxima, tendo como alvo a célula adjacente com menor valor de piso.

1. Para o cálculo da rota, um indivíduo do reticulado observa se sua célula alvo está desocupada, caso esteja ele opta por esperar que ela seja desocupada ou avança para a célula adjacente a da célula alvo. Com isso, o caminho de todos os pedestres é calculado.
2. Tentativas para que os pedestres possam avançar sobre o reticulado são realizadas. Todas as células são observadas e seu endereço é comparado com as rotas calculadas.
 - (a) Se uma rota corresponde ao valor da célula em observação, verifica-se a vizinhança para a constatação de novos candidatos.
 - (b) Se não há mais candidatos o pedestre é movido.
 - (c) Caso o número de candidatos, a ocupar a célula alvo, seja maior que “1”; a disputa é decidida lançando-se um número aleatório, e o candidato vencedor tem direito de mover-se para a célula pretendida e o perdedor recalcula sua rota.
3. Se um indivíduo alcança a porta, ele é eliminado para não impedir os demais de evacuarem.
4. Os passos (1), (2) e (3) são refeitos até que todos os pedestres sejam movidos.

As principais diferenças dos modelos é definida a seguir: no modelo de [Varas et al. 2007] o pedestre que não se movimentou só prosseguiria no próximo passo de tempo Figura 4 (a). No modelo em [Carneiro 2012] o pedestre obrigatoriamente se moveria para a célula com menor valor adjacente da célula anteriormente escolhida Figura 4 (b). Considerando-se a fusão das técnicas propostas nesses dois modelos, o delineamento do modelo apresentado nesse trabalho, estabelece que o pedestre poderá aguardar que sua célula desejada seja liberada com a movimentação do pedestre que a ocupa, ou escolher outra célula para a realização do movimento Figura 4 (c).



Figura 4. Movimentação de dois pedestres em caso de conflito (a) Modelo de [Varas et al. 2007] (b) Modelo de [Carneiro 2012] (c) Modelo proposto.

5. Resultados experimentais

Inicialmente foram realizados alguns testes para a verificação de formação de padrões durante a evacuação que ocorrem no modelo proposto. O principal comportamento observado, e que distingue o modelo apresentado dos demais já relatados na literatura de ACs, é a tendência da formação de filas, resultante das adaptações lógicas do modelo precursor, ou seja, os grupos de indivíduos tendem a deslocar-se ordenadamente em direção às saídas.

A técnica proposta também foi testada em uma estrutura complexa (com obstáculos) similar a uma estrutura real. O novo experimento levou em consideração, locais aleatórios para a inserção dos pedestres. Para os pedestres, 130 células foram reservadas para cada simulação. A planta da área estudada pode ser vista na Figura 5.

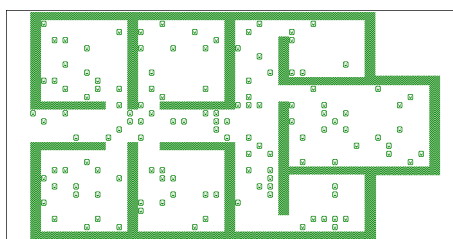


Figura 5. Estrutura complexa simulada a partir de 130 indivíduos gerados aleatoriamente.

O histograma da Figura 6 aponta os dados obtidos com a simulação da planta da Figura 5 a partir de 100 execuções. As classes de intervalos no gráfico representam as quantidades de iterações necessárias para que todos os pedestres pudessem evacuar da estrutura complexa. Dessa forma, pode-se perceber que o modelo de movimentação é bastante eficiente no que diz respeito à distribuição de quantidades de passos de evacuação, uma vez que, o histograma apresenta uma distribuição próxima à normal. Isso significa que os valores obtidos não sofreram grande variação em relação à média ($\bar{X} = 150$ passos). Assim, os resultados gerados pelo modelo foram convergentes para um valor fixo. Considerando-se o tempo de passo 0,4s proposto em [Varas et al. 2007], então o tempo de evacuação médio foi cerca de 60 segundos. Esse tempo é perfeitamente aceitável, embora o local mais distante da estrutura esteja localizado cerca de 20m da saída todos os corredores são traçados com ângulos retos e as barreiras são ortogonais, diminuindo o tempo de evacuação.

Posteriormente, mais experimentos a fim de definir a posição estratégica das portas de saída foram realizados. Para isso, foram adotadas as mesmas características de ambiente testadas por [Varas et al. 2007]. Para facilitar a análise, apenas as estruturas complexas foram utilizadas, como a exibida na Figura 7 (a), ocupada por 50 indivíduos. Além disso, as portas foram inicialmente rotuladas e foram alternadas a cada bloco de testes. Nesse caso, as portas são consideradas separadamente e somente um indivíduo pode retirar-se da sala em um determinado passo de tempo, ou seja, a saída só é permitida por uma única pessoa. Para a análise do local mais indicado para as saídas é observado o rótulo de saída que possui a menor quantidade de iterações. No total, 10 simulações são realizadas, para cada uma das 30 posições “diferentes” das saídas indicadas Figura 7(a). Os resultados observados em [Varas et al. 2007] resultou em médias que variam de 98 a

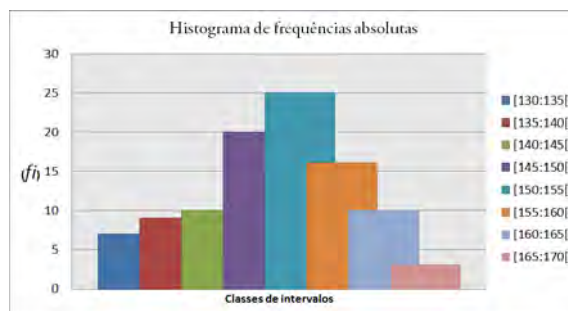


Figura 6. Histograma que representa as frequências de dados obtidas para cada classe (quantidade de passos necessários para a evacuação).

106 passos de tempo para cada porta alternada na mesma estrutura, conforme o gráfico da Figura 8(a). Realizando os testes em estruturas complexas, algumas diferenças surgiram entre os dois modelos. Nesse caso, notou-se que a evacuação é realizada mais rapidamente e as médias variam de 52, 1 a 95 passos de tempo para cada porta modificada, por outro lado, a melhor rota de fuga está próximo à saída 30, conforme também foi observado no modelo de [Varas et al. 2007]. O tempo de evacuação no modelo proposto está abaixo dos valores obtidos por [Varas et al. 2007] e pode ser explicado devido ao fato de que no modelo precursor, apenas um pedestre se movia em caso de conflito de células e o outro ficava paralisado. No modelo aqui descrito, esse pedestre pode escolher uma rota alternativa, além disso, o caso de pânico não foi considerado neste trabalho.

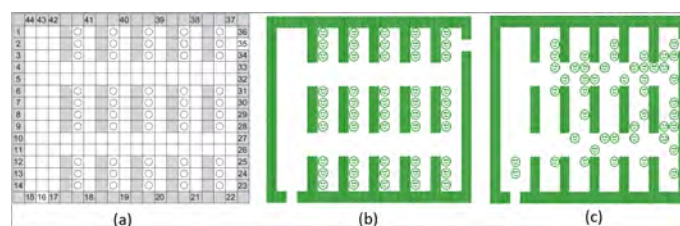


Figura 7. Estrutura: (a) [Varas et al. 2007], (b) geometria aplicada ao modelo proposto e em (c) modelo proposto após 15 iterações.

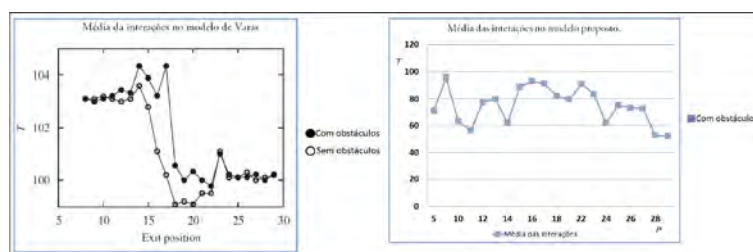


Figura 8. (a) Modelo em [Varas et al. 2007] média das iterações(T) em função da localização da porta na estrutura. (b) Resultados dos testes obtidos com o modelo proposto considerando apenas as estruturas complexas, ou seja, salas com obstáculos.

6. Conclusões e trabalhos futuros

Neste trabalho foi investigado um modelo de autômatos celulares bidimensionais fundamentado em modelos precursores aplicados a simulação de evacuação emergencial para

ambientes complexos. Para a modelagem do AC detalhado no trabalho, considerou-se a implementação descrita por [Varas et al. 2007] e as complementações idealizadas por [Carneiro 2012]. Todas as adaptações lógicas foram concebidas com a análise desses modelos precursores. Além disso, uma análise das características do modelo desenvolvido foi abordada.

Com a execução do modelo descrito foi possível gerar dados visuais, onde os comportamentos de padrões sociais inseridos durante a evacuação puderam ser observados. Além disso, dados estatísticos foram obtidos relacionando a média das iterações com a quantidade de indivíduos alocados em uma estrutura. A grande vantagem de se usar o modelo desenvolvido, na simulação de evacuação emergencial, encontra-se no fato de que ocorre uma padronização das rotas de pedestres. Entretanto, não podemos garantir que apenas os comportamentos inseridos por si só já comprovam a segurança da estrutura predial gerada pelo modelo. Mas, elas iniciam uma análise que pode ser aprofundada em um trabalho de pesquisa posterior.

Como continuidade desse trabalho, seria interessante a adaptação do modelo 2D para um modelo 3D, no qual as alturas dos obstáculos poderiam ser avaliadas. Para alcançar o paralelismo do modelo poderia ser utilizada uma implementação em hardware paralelo, uma vez que toda a motivação para o emprego de ACs na simulação reside na possibilidade de implementação eficiente. Além disso, acredita-se que, o fato do modelo ser gerenciado por regras simples na função de transição, é possível a adição de novos parâmetros para que diferentes comportamentos sociais possam ser observados no decorrer de uma mesma simulação. Essa inserção visa o objetivo do alcance de uma maior fidelidade se comparado com uma situação real.

Referências

- Burstedde, C., Klauck, K., Schadschneider, A., and Zittartz, J. (2001). *Simulation of pedestrian dynamics using a two-dimensional cellular automaton.*, pages 507 – 525.
- Carneiro, L. O. (2012). Simulação de evacuação de multidão por autômato celular - estudo de caso em um estádio de futebol. Master's thesis, Universidade Federal do Ceará.
- Conway, J. H. (1970). Jogo da vida. *Scientific American*. Jogos matematicos.
- Lima, D. A. (2010). Modelos de criptografia baseados no cálculo de pré-imagem de autômatos celulares. *Trabalho de Conclusão de Curso em Ciência da Computação - Universidade Federal de Uberlândia*, pages 20 – 21.
- Oliveira, G. M. B. (2003). Automatos celulares: aspectos dinâmicos e computacionais. *III Jornada de Mini-cursos em Inteligência Artificial (MCIA) - Sociedade Brasileira de Computação*, 8:297 – 345.
- Perez, G., Tapang, G., Lim, M., and Saloma, C. (2002). *Streaming, disruptive interference and power-law behavior in the exit dynamics of confined pedestrians.*, pages 609 – 618.
- Varas, A., Cornejo, M., Mainemer, D., Toledo, B., Rogan, J., Munoz, V., and Valdivia, J. (2007). Cellular automaton model for evacuation process with obstacles. *Physica A: Statistical Mechanics and its Applications*, 382(2):631 – 642.
- Wolfram, S. (2002). *A New Kind of Science*. Wolfram Media - (1st edition): 1197 - 2006-09-19T07:35:05.000+0200.

Uma abordagem de código único para aplicações independentes de provedor de bases de dados relacionais

Willian Eduardo de Moura Casante¹

¹Fundação CPqD – Centro de Pesquisa e Desenvolvimento em Telecomunicações
Rod. SP-340 km 118,5 CEP 13.086-902 – Campinas – SP – Brasil

wcasante@cpqd.com.br

Abstract. *Relational databases uses specific SQL extensions for their product functions, which makes the development of complex applications unfeasible mainly because compatibility with more than one relational database would require specific codes for each database. This paper describes the development of multi-database applications based on a methodology of creation of the installation artifacts for the database and a technique to develop multi-database sources codes.*

Resumo. *Os bancos de dados relacionais fornecem funcionalidades específicas de produtos através de extensões da linguagem SQL, inviabilizando o desenvolvimento de aplicações complexas e compatíveis com mais de um banco de dados relacional sem que essas aplicações tenham código específico para cada um deles. Este artigo aborda o desenvolvimento de aplicações totalmente independentes de banco de dados através de uma metodologia de criação de artefatos de instalação da estrutura de dados e de uma técnica de desenvolvimento de código-fonte independentemente de produto.*

1. Introdução

A linguagem SQL (*Structured Query Language*) [Chamberlin et al. 1981] é uma linguagem de pesquisa e definição de estrutura de dados, inicialmente criada para SGBDs (Sistemas Gerenciadores de Banco de Dados), fortemente inspirada na álgebra relacional [Elmasri et al. 2005].

Necessidades não mapeadas no padrão SQL desenvolvido e aceito pela indústria exigiram dos fabricantes de SGBDs a definição de construções SQL específicas para seus produtos e a criação de extensões da linguagem que possibilitam o acesso das suas funcionalidades proprietárias.

Essas diferenças entre as diversas implementações da linguagem SQL prejudicam o desenvolvimento de um sistema computacional compatível com dois ou mais SGBDs pois, para cada SGBD atendido, um conjunto de arquivos específicos para aquele deverá ser mantido, tanto para a instalação do sistema quanto para o funcionamento correto de seu código-fonte.

2. Análise da incompatibilidade entre SGBDs

Um sistema computacional que realiza o armazenamento de suas informações em SGBDs possui seus arquivos de instalação organizados em DDLs (*Data Definition Language*) e DMLs (*Data Manipulation Language*).

DDL é um arquivo de metadados, que define a estrutura das tabelas de um banco de dados, composto por comandos SQL como, por exemplo, *CREATE*, *ALTER* e *DROP*. Os arquivos do tipo DDL são os primeiros a serem executados durante a instalação de um sistema computacional, visto que criam o arcabouço do sistema no SGBD.

DML é um arquivo de dados, cuja estrutura respeita fielmente a estrutura definida em seu respectivo arquivo DDL. Durante o desenvolvimento de um sistema pode, se necessário, inicializar o banco de dados; essa inicialização é colocada em arquivos do tipo DML. Os arquivos DML possuem comandos de *INSERT*, *UPDATE*, *DELETE* e *SELECT*, também utilizados de forma embutida dentro do código-fonte do sistema.

2.1. A estrutura dos dados em um SGBD

Durante a análise de um sistema, uma das tarefas é a definição de seu modelo de banco de dados. Essa definição normalmente envolve a construção de entidades básicas do sistema, em uma linguagem visual de alto nível, como, por exemplo, o modelo de entidades e relacionamentos [Barbieri 1994] ou a UML (*Unified Modeling Language*) [Jacobson et al. 1996]. Um exemplo simples de modelo UML pode ser observado na Figura 1. Através desse modelo inicial, os analistas de sistemas realizam um aprofundamento na modelagem, através da definição de todas as entidades e de seus relacionamentos, para que o sistema possa ser construído sobre esse modelo proposto. Conforme o sistema é desenvolvido, novas entidades e relacionamentos são incluídos e estes, detalhados e então o sistema evoluído para os contemplar.

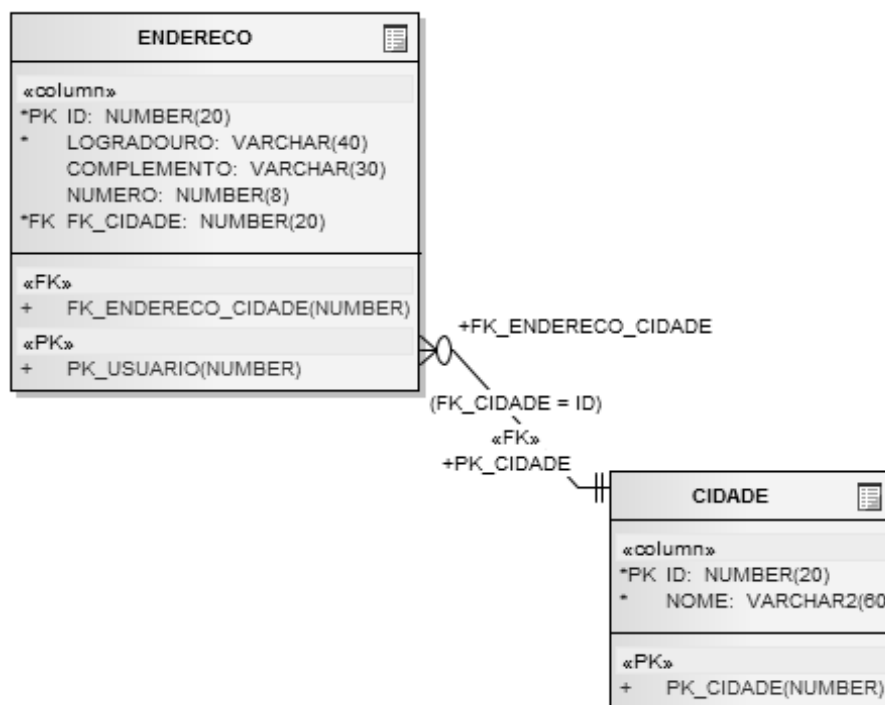


Figura 1. Exemplo de modelagem via UML

A partir da definição gráfica das entidades do modelo de banco de dados, um especialista deverá gerar um ou mais arquivos DDL, contendo os comandos SQL equivalentes

para a criação daquele modelo em um SGBD específico. Existem ferramentas que geram automaticamente esses arquivos a partir dos diagramas propostos, que estão fora do escopo deste artigo.

Um sistema que se propõe a ser instalado em mais de um SGBD apresenta algumas estratégias que podem ser adotadas, para a criação de seus arquivos DDL:

- Criação de um ou mais arquivos DDL por SGBD – exige a replicação de artefatos de trabalho (arquivos DDL) para cada um dos SGBDs; adiciona risco de erro na conversão, caso realizada de forma manual, como o risco da falta de sincronia entre as atualizações dos artefatos. Como exemplo, um profissional obtém uma DDL para o SGBD Oracle e deve convertê-la para funcionar no SGBD PostgreSQL e causa um erro ao não trocar NUMBER(20) por NUMERIC(20);
- Criação de um único arquivo DDL para um SGBD e utilização de uma ferramenta de conversão para os outros SGBDs – a conversão por ferramenta mitiga os possíveis erros porém adiciona um passo à geração do pacote de instalação do sistema: a conversão dos arquivos DDL. A existência de uma ferramenta que realize essa conversão, bem como sua aquisição, caso seja comercial, é um requisito primário, que pode ser substituído pelo desenvolvimento da própria ferramenta, como realizado no desenvolvimento da ferramenta JExodus [Neto and Passos];
- Geração automática de modelo através de código-fonte – alguns arcabouços de programação (por exemplo a especificação EJB (*Entity JavaBean*) do Java EE [Patel et al. 2006]) permitem a definição de entidades programáticas e a geração do modelo de banco de dados de forma transparente. Esse formato não é aproveitável em sistemas que necessitam controlar com precisão seus modelos de banco de dados ou realizar atualizações pois a instalação e a definição do modelo ficam a encargo do arcabouço. Pode haver conflito no que diz respeito à segurança, onde a instalação da estrutura do banco de dados ficar a cargo de um administrador do SGBD, ou durante a execução, se o sistema obtiver uma conexão na qual apenas controla a atualização dos dados neste SGBD.

2.2. A manutenção dos dados em um SGBD

A definição da estrutura, ou modelo, de um sistema em um SGBD, conforme descrito na Subseção 2.1, deverá permitir a inserção, alteração, exclusão e consulta dos dados presentes nessa estrutura. Para realizar essas operações existem os comandos DML.

A instalação de um sistema pode necessitar de dados iniciais, inseridos durante a própria instalação ou configurados por uma equipe de implantação, dados estes presentes em scripts DML e também, durante a utilização do sistema, realiza esses comandos no SGBD para que os dados sejam manipulados na citada estrutura.

Portanto, podemos listar dois cenários a serem tratados:

- A criação dos dados iniciais em uma estrutura, durante a instalação de um sistema, cujos comandos DML deverão ser compatíveis com diversos SGBDs ou convertidos conforme necessidade;
- O ciclo de produção do sistema, pós-instalação, no qual o sistema emitirá comandos DML para a manipulação dos dados presentes na estrutura instalada. Esses comandos deverão ser genéricos ou estruturados de forma que o sistema não

conheça as diferenças entre os diversos SGBDs e suas peculiaridades, permitindo que o código-fonte, geralmente em linguagem de alto nível (como o Java), seja único e independente de SGBD.

2.3. Linguagens específicas

Cada SGBD pode prover uma ou mais linguagens de programação específicas, conhecidas como linguagem de *script*, nas quais é possível adicionar lógica de programação além dos comandos DDL e DML. Essas linguagens (por exemplo, o PL/SQL do SGBD Oracle [Feuerstein 2002]) possuem sintaxe proprietária e muitas vezes de difícil conversão automática entre os diversos SGBDs.

Podem ser definidas funções (*Stored Procedures*), diretamente no SGBD, utilizando-se essas linguagens de programação e realizar a chamada dessas funções diretamente do código-fonte do sistema.

3. Arquiteturas passíveis de utilização

Enumeram-se algumas possíveis arquiteturas de acesso e manutenção dos dados de múltiplos SGBDs, para um sistema de software:

- Desenvolvimento tradicional – com projeto de banco de dados, geração de arquivos DDL e DML, controle manual das alterações e controle de arquivos por SGBD;
- Arcabouço de desenvolvimento, como o Ruby on Rails [St.Laurent et al. 2012] ou o Java Persistence API [Patel et al. 2006], nos quais o próprio arcabouço trata a criação e a evolução do banco de dados;
- Metodologia proposta neste artigo, que unifica os itens citados anteriormente na manutenção de uma única base de artefatos relativos à base de dados.

4. Adotando uma metodologia única

A metodologia, proposta na Figura 2, une o desenvolvimento tradicional aos arcabouços de desenvolvimento, aproveitando o seu melhor sem limitar-se às fronteiras impostas pelos arcabouços, facilitando assim o uso em sistemas de grande porte, que exigem a modelagem de bases de dados extremamente complexas.

A premissa para o desenvolvimento do sistema é a capacidade de este ser executado em dois ou mais SGBDs. Se essa necessidade não existir, o desenvolvimento tradicional poderá ser adotado. Do mesmo modo, se a complexidade da estrutura de dados for baixa, independentemente do tamanho do sistema, será possível utilizar por completo e com todas as vantagens, um dos arcabouços de desenvolvimento ágil (como, por exemplo, o Ruby on Rails).

4.1. Definição de um SGBD “base”

Antes de iniciar o desenvolvimento, para determinar o SGBD “base”, no qual os arquivos DML e DDL serão escritos, deverão ser realizados uma pesquisa e um estudo das ferramentas disponíveis para conversão de cada SGBD suportado pelo sistema. Muitas ferramentas requerem de licença de uso, o que deve ser levado em conta na análise e na tomada de decisão. Pode-se chegar à conclusão de que não há ferramenta compatível com os requisitos impostos, e, nesse caso, será utilizada a conversão manual. É recomendável adicionar ao projeto o risco de defeitos causados pela conversão manual dos artefatos.

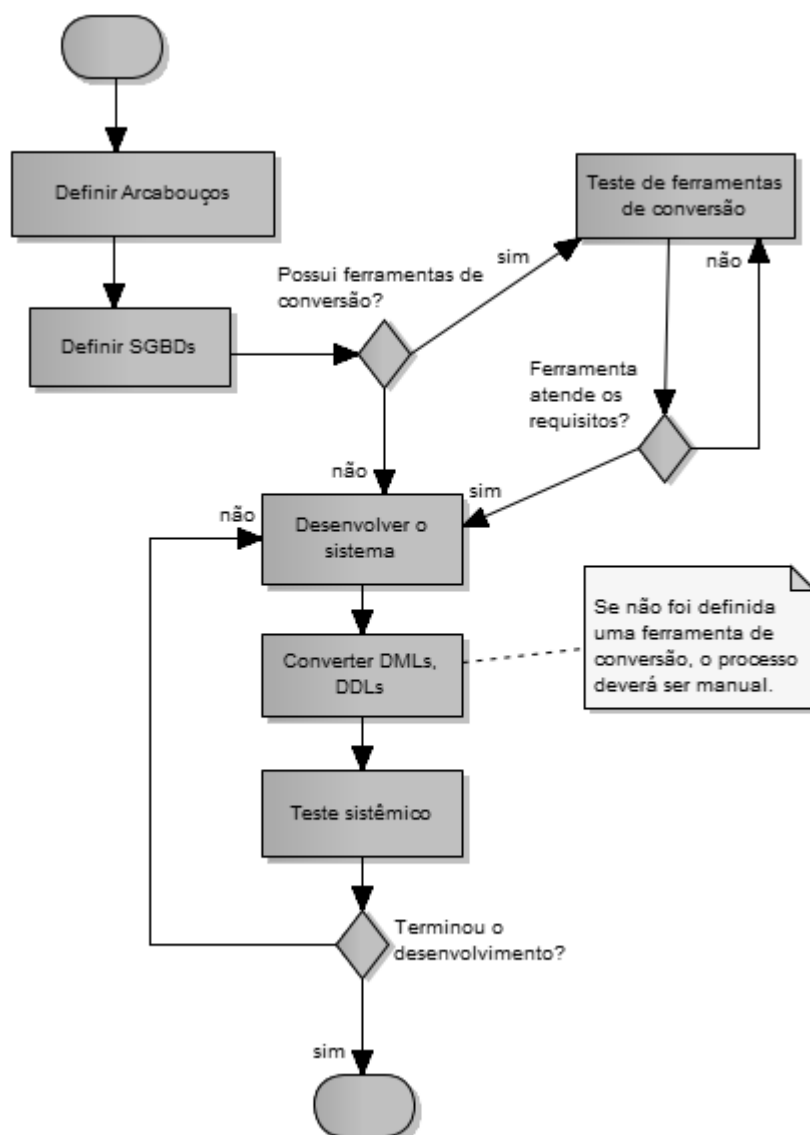


Figura 2. Metodologia sugerida para o desenvolvimento do sistema

4.2. Definição dos arcabouços de programação

Uma vez conhecida a linguagem de programação na qual o sistema será desenvolvido, aconselha-se escolher também um arcabouço de abstração de acesso aos dados. Esses arcabouços abstraem detalhes dos diversos SGBDs suportados, padronizam o código-fonte, e facilitam o desenvolvimento do acesso aos dados, pois normalmente oferecem facilidades para tal.

Deve ser feito um estudo de uso, conforme feito por Viana e Borba [Viana and Borba 1999], a partir de arcabouços atuais da linguagem de programação escolhida (por exemplo, para o Java, o Hibernate [Bauer and King 2005] ou a especificação Java Persistence API [Patel et al. 2006]), para determinar o melhor conjunto de arcabouços. Caso os SGBDs escolhidos suportem a orientação à objetos [Boscarioli et al. 2006], deve-se observar se os arcabouços de programação escolhidos possuem o suporte necessário para as funcionalidades que serão utilizadas.

4.3. Ciclo de vida do sistema

Definidos o SGBD “base” e a ferramenta de conversão para os outros SGBDs (se for o caso), são gerados os scripts DML e DDL e realizadas as conversões necessárias, conforme o sistema é desenvolvido. Eventuais scripts de atualização de versão do sistema precisarão ser escritos para cada SGBD, pois muitas vezes envolvem migração de dados.

O uso de funções SQL diretamente em banco de dados deve ser evitado, pois aumenta a complexidade da solução, exigindo sua manutenção para cada SGBD, bem como a coerência entre as diferentes codificações, inerentes a cada SGBD.

É importante que todos os envolvidos no desenvolvimento do sistema sigam o processo a partir das soluções adotadas, tanto para a conversão entre os diversos SGBDs quanto para o desenvolvimento do próprio sistema na sua linguagem de programação. Esse cuidado garante que todos trabalhem para que a aplicação continue compatível com cada SGBD escolhido.

O teste sistêmico deverá ser realizado com o sistema instalado em cada SGBD, garantindo assim que esteja compatível com todos.

É importante entender que, quanto maior o nível de abstração oferecido pelo arcabouço escolhido e de automação na conversão entre os SGBDs menor o risco de incompatibilidade entre eles. Para um sistema totalmente abstraído em relação ao seu modelo relacional, o teste sistêmico poderá ser realizado por amostragem, sendo que os casos de teste são executados parcialmente em um SGBD e parcialmente em outro, evitando que seja necessário realizar todos os casos de teste em todos os SGBDs, com expressivo ganho de tempo durante seu desenvolvimento.

5. Prova de conceito

Para a realização da prova de conceito foi considerado um sistema inicialmente construído para o SGBD Oracle, que deveria ser compatível também com o SGBD PostgreSQL. O requisito de “portabilidade” entre ambos SGBDs, durante a vida do sistema e suas evoluções, possibilitou a aplicação da metodologia definida na Seção 4, garantindo a qualidade dos artefatos gerados e a ausência de impacto nos prazos de entrega, eliminando também o risco produzido por conversões manuais realizadas pelos desenvolvedores.

A estrutura adotada, para os artefatos gerados para o projeto é apresentada na Figura 3.

Conforme apresentado na Seção 2, existem arquivos DDL, de instalação da estrutura de tabelas do sistema e arquivos DML, que realizam a população dos dados iniciais necessários para o funcionamento do sistema, bem como os comandos DML presentes em código-fonte, no caso, escrito em Java.

Os comandos DDL, originalmente em Oracle, foram convertidos utilizando-se a ferramenta Ora2Pg [Ora2Pg 2012]. Esses arquivos precisam inicialmente ser instalados em um SGBD Oracle para que a ferramenta consiga extrair as DDLs no formato PostgreSQL. Da mesma forma que as DDLs, as DMLs de carga do sistema foram instaladas no banco de dados Oracle e extraídas com a mesma ferramenta para o formato PostgreSQL.

Os arquivos DDL e DML gerados para o PostgreSQL são colocados em controle

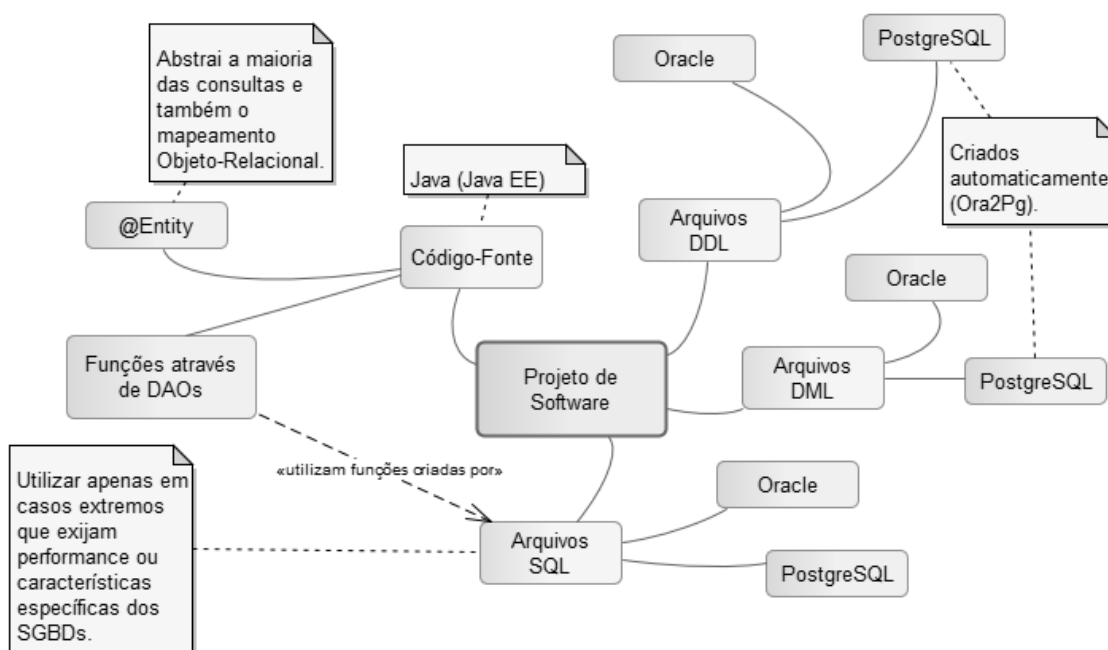


Figura 3. Estrutura do projeto da prova de conceito

de versão, sendo disponibilizados para a instalação, juntamente com o produto. Dessa forma o instalador poderá optar por fazer a instalação em um SGBD Oracle ou PostgreSQL, utilizando para isso um único pacote de instalação.

Funções SQL podem ser criadas diretamente em banco de dados e poderão ser utilizadas, a partir do código-fonte Java, através do padrão de projeto DAO (*Data Access Object*) [Alur et al. 2003]. Essas funções devem ser criadas com uma mesma assinatura de chamada, mesmo que para diferentes SGBDs, para que a chamada possa ser única e a construção do DAO independente do SGBD, bem como devem ser construídas para todos os SGBDs definidos.

O sistema foi construído totalmente sobre o arcabouço Java EE EJB e todas as entidades de banco de dados foram mapeadas utilizando-se as funcionalidades de abstração fornecidas por esse arcabouço. Todas as consultas a essas entidades foram construídas utilizando-se a linguagem de consulta JPQL (*Java Persistence Query Language*), que abstrai a linguagem SQL, garantindo que o código-fonte Java seja único e independente do SGBD adotado.

O resultado da prova de conceito é um sistema capaz de trabalhar tanto no SGBD Oracle quanto no SGBD PostgreSQL, com um único pacote de instalação e um único código-fonte, sendo que os processos de conversão entre esses SGBDs são completamente automatizados, validando a metodologia proposta.

6. Conclusão

Propor uma solução para a abstração de acesso aos dados de um SGBD, bem como centralizar e automatizar a geração de comandos de definição do modelo relacional, garante a longevidade do sistema desenvolvido. O sistema torna-se independente de SGBD, permitindo a redução dos custos de implantação pela adoção de soluções de software livre

sem adaptações do código-fonte original ou mesmo troca estratégica do SGBD. O conhecimento dos desenvolvedores envolvidos na produção do sistema pode ser nivelado, pois não são mais necessários especialistas para cada SGBD, os riscos são mitigados e os pontos passíveis de erro reduzidos.

Por fim, a automação da geração e do controle dos artefatos de cada SGBD evita erros relacionados à migração de código-fonte específico de determinado SGBD, que são maiores no caso de migração manual, bem como reduz o custo do desenvolvimento por eliminar as horas necessárias para a conversão.

Referências

- Alur, D., Crupi, J., and Malks, D. (2003). *Core J2EE patterns: best practices and design strategies*. Prentice Hall.
- Barbieri, C. (1994). *Modelagem de Dados*. IBPI Press Rio de Janeiro.
- Bauer, C. and King, G. (2005). *Hibernate in action*. Manning.
- Boscarioli, C., Bezerra, A., Benedicto, M., and Delmiro, G. (2006). Uma reflexão sobre banco de dados orientados a objetos. *IV CONGED*.
- Chamberlin, D., Astrahan, M., Blasgen, M., Gray, J., King, W., Lindsay, B., Lorie, R., Mehl, J., Price, T., Putzolu, F., et al. (1981). A history and evaluation of system r. *Communications of the ACM*, 24(10):632–646.
- Elmasri, R., Navathe, S., Pinheiro, M., Canhette, C., Melo, G., Amadeu, C., and de Oliveira Morais, R. (2005). *Sistemas de banco de dados*. Pearson Addison Wesley.
- Feuerstein, S. (2002). *Oracle pl/sql Programming*. O'Reilly.
- Jacobson, I., Booch, G., and Rumbaugh, J. (1996). *The Unified Modeling Language*. University Video Communications.
- Neto, J. and Passos, E. Jexodus: uma ferramenta para migração de dados independente de sgbd.
- Ora2Pg (2012). Ora2pg software. <http://ora2pg.darold.net/>.
- Patel, R., Brose, G., and Silverman, M. (2006). *Mastering Enterprise JavaBeans 3.0*. Wiley Pub.
- St.Laurent, S., Dumbill, E., and Gruber, E. (2012). *Learning Rails 3*. O'Reilly.
- Viana, E. and Borba, P. (1999). Integrando java com bancos de dados relacionais. *III Simpósio Brasileiro de Linguagens de Programação, Porto Alegre Brasil*, 5.

Metodologia de análise para migração de sistemas de médio e grande portes para a nuvem no modelo SaaS

Cláudia Morgado^{1,2}, André Temple de Antonio¹, Rodrigo Lima Verde Leal¹,
Regina Lúcia de Oliveira Moraes²

¹Fundação CPqD – Centro de Pesquisa e Desenvolvimento em Telecomunicações
R. Dr. Ricardo Benetton Martins, s/n – Parque II do Polo de Alta Tecnologia – CEP
13.086-902 – Campinas – SP – Brasil

²Faculdade de Tecnologia
Universidade Estadual de Campinas – Unicamp – Limeira – SP – Brasil
{cmorgado, andret, rodleal}@cpqd.com.br, regina@ft.unicamp.br

Abstract. *Even though cloud computing provides on-demand access to easily activated and practically unlimited IT resources on a pay-as-you-use model, there are still hindering factors that should be considered before migrating a system to the cloud. Each system has its own particularities that must be individually analyzed in order to assess the impacts of putting a system in the cloud. This paper proposes a methodology to analyze the characteristics of large systems that must be considered in the migration to the cloud.*

Resumo. *Apesar de a computação em nuvem fornecer acesso a recursos de tecnologia da informação sob demanda, praticamente ilimitados, facilmente ativados e com modelo de pagamento baseado no uso, existem diversos fatores restritivos que devem ser analisados antes de se optar pela projeção ou migração de um sistema para a nuvem. Desse modo, as particularidades de cada sistema devem ser analisadas individualmente, a fim de avaliar o impacto e a viabilidade de se disponibilizar o sistema na nuvem. Este artigo propõe uma metodologia para avaliação das características de sistemas de grande porte que devem ser consideradas no processo de migração para a nuvem no modelo SaaS.*

1. Introdução

O modelo da computação tradicional possui uma necessidade de infraestrutura de TI complexa. É necessário que a empresa disponha de um ambiente computacional com servidores e computadores, de locais especiais de armazenamento com climatização adequada e de uma equipe de especialistas sempre prontos para realizar eventuais manutenções, o que eleva consideravelmente os custos. Para agravar ainda mais esse cenário, com o passar do tempo, esses servidores se tornam obsoletos, sendo necessária sua substituição por máquinas mais robustas, o que também gera altos custos.

No conceito de computação em nuvem (CN), não é mais necessário que uma empresa possua um servidor e diversas máquinas de alto desempenho dentro do ambiente de trabalho. Basta apenas que a empresa contrate um provedor de serviços de CN para ter acesso a infraestrutura, plataforma ou software, de uma forma simples e a

partir de qualquer local, por meio da Internet, sem que seja necessário instalar, configurar ou atualizar os sistemas de software.

Segundo o NIST (National Institute of Standards and Technology), "Computação em nuvem é um modelo que possibilita acesso, de modo conveniente e sob demanda, a um conjunto de recursos computacionais configuráveis que podem ser rapidamente adquiridos e liberados com mínimo esforço gerencial ou interação com o provedor de serviços" [NIST 2012 apud Hogan et al. 2011].

Redução de gastos, maior agilidade e flexibilidade são fatores que levam as empresas a disponibilizarem seus sistemas na nuvem. Algumas estimativas apontam que, em dois ou três anos, pelo menos 80% das empresas de software terão adotado ou estarão em processo de adoção desse modelo [Taurion 2009]. Contudo, a migração de sistemas atuais para um ambiente de computação em nuvem traz uma série de desafios que tendem a dificultar a adoção desse novo paradigma por muitos usuários e organizações.

Neste artigo apresentamos uma metodologia de apoio aos tomadores de decisão que os ajudará a compreender as características da nuvem e identificar se um software desenvolvido na computação tradicional tem potencial para ser disponibilizado como serviço. A disponibilização de um software como serviço é conhecida como o modelo de serviço na CN SaaS. No modelo SaaS, o usuário não administra ou controla a infraestrutura subjacente, que inclui rede, servidores, sistema operacional, armazenamento, ou mesmo as características individuais da aplicação, exceto as configurações específicas [Sousa, Moreira, Machado 2010].

Essa metodologia surgiu em decorrência da necessidade de avaliação da viabilidade da migração de sistemas para a nuvem.

2. Impactos na evolução de um sistema de software para a nuvem

As vendas no modelo SaaS são baseadas no serviço sob demanda, ou seja, o cliente paga apenas o que utilizar e isso, na prática, impacta as áreas de marketing e vendas de uma empresa, que devem estar preparadas para medir o uso do sistema [Taurion 2009]. O sistema também deve estar preparado para atender às necessidades do cliente à medida que elas surgem, como, por exemplo, tempo de uso do sistema e armazenamento de dados, sem requerer interação humana com o cliente [Sousa, Moreira, Machado 2010].

O software pode ser acessado via Web com interface de acesso que não obrigue os usuários a mudarem suas condições e seu ambiente de trabalho. As modificações devem ser apresentadas de forma transparente [Jansen e Grance 2011].

Como a proposta do modelo SaaS é ter uma única aplicação que atenda a múltiplos clientes (empresas), chamados *tenants* ou inquilinos, a arquitetura essencial de um ambiente na nuvem é do tipo *multi-inquilino*, uma vez que permite que múltiplos inquilinos compartilhem os mesmos recursos físicos e permaneçam logicamente isolados [Cambiucci 2009].

3. Metodologia para avaliação de software

O controle de um sistema disponibilizado na nuvem é extremamente complexo se comparado ao sistema tradicional e impacta áreas distintas da empresa. Por essa razão,

se faz necessária uma metodologia para auxiliar no processo de decisão estratégica da empresa que identifique os pontos da aplicação que podem ter maior impacto durante uma migração para a nuvem. Desse modo, a empresa pode avaliar se a aplicação será totalmente migrada para a nuvem, se apenas parte do sistema deve ser migrada ou, ainda, se a migração é uma nova oportunidade de negócio para a empresa. [Taurion 2009].

A metodologia propõe a aplicação de um questionário composto de 26 perguntas que abrangem quatro modalidades do sistema a ser migrado. Questões sobre a própria aplicação a ser migrada, análise das regras de negócio em que a aplicação está inserida, tecnologia usada no desenvolvimento e características de nuvem.

Cada questão recebe uma pontuação que indica o impacto durante a migração, sendo que quanto maior a pontuação, maior o risco de uma migração para a nuvem.

O questionário deve ser respondido por um profissional que conheça os aspectos de marketing do produto e a arquitetura do sistema. As questões estão baseadas nas características de nuvem levantadas pelo NIST [NIST 2012 apud Mell e Grace 2011], em relatos de experiência com migração de sistema para nuvem [Chauhan e Babar 2011], e procura esclarecer os pontos em que o sistema escolhido para análise de viabilidade já trabalha no conceito de nuvem e os pontos em que a migração traria um maior impacto.

A ferramenta de apoio à decisão na adoção para nuvem proposta pela Universidade de St. Andrews propõe a aplicação de um *checklist* para a análise tecnológica de um sistema [Hosseini et al 2012]. Mediante a utilização desse *checklist* em nossos produtos, foi verificada a necessidade de incluir novas questões e diferenciar o peso das questões com base no impacto que cada uma delas traz ao processo de migração.

A **Tabela 1** mostra as questões elaboradas com o objetivo de se obter um melhor entendimento da aplicação, como, por exemplo, nicho de mercado e características do software que podem dificultar a migração para nuvem.

Tabela 1. Análise da aplicação a ser migrada

Questionário	Justificativa	Pontuação	
		Sim	Não
1. É uma solução para um nicho específico de clientes?	<i>Analisar se a solução da aplicação foi feita especificamente para um único cliente ou se ela pode atender a um nicho de mercado com problemas e necessidades bem definidos..</i>	0	5
2. Esta aplicação retém seus usuários?	<i>Avaliar se a aplicação pode ser facilmente substituída por outra.</i>	0	5
3. Esta aplicação de software conecta o negócio do contratante com seus fornecedores, empregados, investidores e clientes?	<i>Analisar se o software suporta a integração com outros módulos, como, por exemplo, faturamento, transferência de fundos, gestão de inventário e CRM.</i>	5	0
4. Esta aplicação ajuda equipes a trabalharem em conjunto, dentro ou entre organizações?	<i>Analisar se o software é utilizado de maneira compartilhada, como, por exemplo, sistemas de calendário, e-mail, compartilhamento de telas, autoria compartilhada de conteúdo, gestão de conferência, jogo on-line. Características de aplicações na nuvem.</i>	0	5
5. Esta aplicação é voltada ao ambiente de escritório?	<i>Processadores de texto, planilhas de cálculo, programas de apresentação ou banco de dados. Essas aplicações possuem características nativas para nuvem.</i>	0	5
6. Esta aplicação resolve	<i>Ferramentas de conversão de formato de arquivo,</i>	0	5

problemas de segurança ou compatibilidade e suporta desenvolvimento de novo software?	<i>digitalização e análise de segurança, verificação de regras e desenvolvimento. Essas aplicações possuem características nativas para nuvem.</i>		
7. Esta aplicação requer tempo de execução de tarefas definido?	<i>Não suporta tempo de espera ou atrasos de entrega.</i>	5	0
8. Esta aplicação requer transferência de grande volume de dados originados no assinante?	<i>Possui integração com outras aplicações com grandes transferências de dados entre elas.</i>	5	0
9. É um sistema de missão crítica?	<i>Uma falha desta aplicação (erro de execução ou lentidão) pode causar perda significativa de dados?</i>	5	0
10. Esta aplicação trabalha com altas taxas de refresh na tela do assinante?	<i>A tela é atualizada continuamente.</i>	5	0

A adoção do modelo SaaS é uma decisão estratégica que afeta o modelo de negócio da empresa [Taurion 2009] e, por esse motivo, a **Tabela 2** apresenta questões para um melhor entendimento do contexto do negócio em que a aplicação está inserida. As questões das tabelas a seguir recebem uma **pontuação (de 0 a 5)**, considerando que quanto maior a pontuação, maior o impacto durante a migração.

Tabela 2. Análise de negócio

Questionário	Justificativa
11. Esta aplicação já possui um modelo de negócio para sua venda como serviço sob demanda, baseado em assinatura (pay-as-you-go)?	<i>Analisar como é feita a venda do software escolhido para evoluir para nuvem, verificar se o produto já é cobrado conforme o uso do sistema e qual seria o grau de impacto de uma possível alteração.</i>
12. Esta aplicação pode ser considerada uma solução em potencial para o mercado de pequenas e médias empresas?	<i>Analisar se ocorreria um crescimento significativo de clientes após a evolução para o modelo SaaS, verificar se o sistema de software passaria a atender a um maior nicho de mercado, incluindo pequenas e médias empresas como clientes do software.</i>
13. Esta aplicação tem uma demanda capilarizada, com grande número de clientes potenciais?	<i>Analisar se a aplicação possui uma quantidade expressiva de clientes que sustentem a disponibilização dos serviços na nuvem.</i>
14. Esta aplicação já possui um modelo contratual de prestação de serviço baseado em SLA (Service Level Agreement)?	<i>O produto é padronizado a todos os clientes e existe um acordo-padrão, firmado entre a área de TI e o cliente, que descreve o serviço de TI, suas metas, papéis e as responsabilidades das partes envolvidas.</i>
15. O retorno financeiro desta aplicação pode ser quantificado?	<i>Verificar se existe algum aplicativo que mostre ao cliente a análise de retorno sobre o investimento (ROI), comparando-a com a da solução tradicional.</i>
16. Existe alguma infraestrutura para hospedagem do software ou alguma parceria a esse respeito?	<i>Avaliar se a empresa já possui infraestrutura ou mantém parceria para hospedar o sistema na nuvem, a partir de um dos modelos possíveis (privada, pública, comunidade ou híbrida).</i>

A **Tabela 3** representa a parte do questionário que tem por objetivo fazer uma análise das características essenciais da CN. Essas características são relacionadas às vantagens que a CN oferece [Sousa, Moreira, Machado 2010].

Tabela 3: Características de nuvem

Questionário	Justificativa
17. A aplicação possui o conceito de self-service sob demanda?	<i>O sistema também deve estar preparado para atender às necessidades do cliente à medida que elas surgem, como, por exemplo, tempo de uso de servidor, armazenamento na rede, sem requerer interação humana com o prestador do serviço.</i>
18. O acesso à aplicação pode ser feito por meio da Internet?	<i>Os recursos estão disponíveis na rede e são acessados por meio de mecanismos padronizados que promovem o uso com plataformas do tipo thin clients heterogêneas (ex.: celulares, tablets, laptops e estações de trabalho).</i>
19. A aplicação permite que múltiplos clientes compartilhem os mesmos recursos, mas permaneçam logicamente isolados?	<i>Na CN os recursos computacionais do prestador de serviço são agrupados (pooling) para servir múltiplos consumidores por meio do modelo multi-inquilino (multi-tenant), como, por exemplo, diferentes recursos físicos e virtuais alocados e realocados dinamicamente, de acordo com a demanda do consumidor. Há um senso de independência da localização, no sentido de que o consumidor geralmente não tem controle ou conhecimento sobre a localização exata dos recursos do prestador de serviço, mas consegue especificar a localização em um nível mais alto de abstração (ex.: país, estado, etc).</i>
20. Os recursos podem ser provisionados e liberados rápida e elasticamente (em alguns casos, automaticamente), para escalar rapidamente dentro e fora de acordo com a demanda?	<i>Para o consumidor, os recursos disponíveis para provisionamento geralmente dão a sensação de ser ilimitados e passíveis de contratação em qualquer quantidade e a qualquer instante.</i>
21. O uso de recursos pode ser monitorado, controlado e reportado, dando transparência tanto para o prestador de serviço quanto para o consumidor do serviço utilizado?	<i>O uso dos recursos é otimizado e controlado automaticamente por meio de capacidades de medição em um nível de abstração apropriado para o tipo de serviço prestado (ex.: armazenamento, processamento, banda e contas de usuário ativas).</i>

A Tabela 4 representa a parte do questionário que tem por objetivo fazer uma análise das características tecnológicas do sistema, tais como: arquitetura, interface e segurança.

Tabela 4. Características tecnológicas

Questionário	Justificativa
22. Esta aplicação possui arquitetura orientada a serviços (Service-Oriented Architecture – SOA)?	<i>SOA é uma arquitetura conceitual de negócios que visa maximizar a reutilização de serviços existentes e a integração com novos serviços para aumentar a eficiência e a produtividade de uma empresa. A aplicação foi desenvolvida com essa metodologia.</i>
23. Esta aplicação é capaz de suportar modelos pré-configurados?	<i>Trata-se de um desdobramento da questão anterior. Todas as alterações devem atender a todos os clientes, permitindo no máximo alguma customização específica de cliente em modelos pré-configurados. No modelo SaaS, deve-se evitar diferentes versões de software para cada cliente.</i>
24. Esta aplicação possui bom desempenho em termos de tempo de resposta?	<i>Todas as funcionalidades possuem respostas imediatas à solicitação do usuário.</i>
25. Esta aplicação oferece alta disponibilidade?	<i>A aplicação fica disponível 24 horas por dia, 7 dias por semana.</i>
26. Esta aplicação possui requisitos de segurança?	<i>A aplicação possui controle de acesso, criptografias, perfil de usuário.</i>

4. Estudo de caso

Os analistas de marketing do produto e arquitetura do software do sistema CPqD Billing responderam o questionário proposto pela metodologia.

A **Tabela 5** mostra as respostas da análise de um sistema de *billing* desenvolvido pela Fundação CPqD – Centro de Pesquisa e Desenvolvimento em Telecomunicações, que atende empresas operadoras de telecomunicações e integra em uma única solução os processos de faturamento, arrecadação, cobrança e atendimento ao cliente.

Tabela 5. Análise do software de *billing* com o uso da metodologia aplicada

Análise da aplicação				Sim	Não	
1. É uma solução vertical, que fornece uma solução específica para um problema ou uma solução para um nicho específico de clientes?					X	
2. Esta aplicação retém seus usuários?					X	
3. Esta aplicação de software conecta o negócio do contratante com seus fornecedores, empregados, investidores e clientes?					X	
4. Esta aplicação ajuda equipes a trabalharem em conjunto, dentro ou entre organizações?					X	
5. Esta aplicação é voltada ao ambiente de escritório?					X	
6. Esta aplicação resolve problemas de segurança ou compatibilidade e suporta desenvolvimento de novo software?					X	
7. Esta aplicação requer tempo de execução de tarefas definido?				X		
8. Esta aplicação requer transferência de grande volume de dados originados no assinante?				X		
9. Uma falha desta aplicação (erro de execução ou lentidão) pode causar perda significativa de dados?					X	
10. Esta aplicação trabalha com altas taxas de <i>refresh</i> na tela do assinante?					X	
Análise de negócio	0	1	2	3	4	5
11. Esta aplicação já possui um modelo de negócio para sua venda como serviço sob demanda, baseado em assinatura (<i>pay-as-you-go</i>)?					X	
12. Esta aplicação pode ser considerada uma solução em potencial para o mercado de pequenas e médias empresas?					X	
13. Esta aplicação tem uma demanda capilarizada, com grande número de clientes potenciais?				X		
14. Esta aplicação já possui um modelo contratual de prestação de serviço baseado em SLA (Service Level Agreement)?			X			
15. O retorno financeiro desta aplicação pode ser quantificado?			X			
16. Existe alguma infraestrutura para hospedagem do software ou alguma parceria a esse respeito?				X		
Análise das características de nuvem	0	1	2	3	4	5
17. A aplicação possui o conceito de <i>self-service</i> sob demanda?					X	
18. O acesso à aplicação pode ser feito por meio da Internet?		X				
19. A aplicação permite que múltiplos clientes compartilhem os mesmos recursos, mas permaneçam logicamente isolados?						X
20. Os recursos podem ser provisionados e liberados rápida e elasticamente (em alguns casos, automaticamente), para escalar rapidamente dentro e fora de acordo com a demanda?						X
21. O uso de recursos pode ser monitorado, controlado e reportado, dando transparência tanto para o prestador de serviço quanto para o consumidor do serviço utilizado?					X	
Análise das características tecnológicas	0	1	2	3	4	5

22. Esta aplicação possui arquitetura orientada a serviços (Service-Oriented Architecture – SOA)?				X		
23. Esta aplicação é capaz de suportar modelos pré-configurados?						X
24. Esta aplicação possui bom desempenho em termos de tempo de resposta?					X	
25. Esta aplicação oferece alta disponibilidade?			X			
26. Esta aplicação possui requisitos de segurança?				X		

As respostas foram pontuadas de acordo com a metodologia. As pontuações foram somadas por etapa, considerando-se fortemente desfavorável a uma migração o valor máximo possível para soma e fortemente favorável para migração para a nuvem a menor pontuação possível. Na Figura 1 os marcadores em negrito representam a somatória da pontuação de cada uma das questões do sistema avaliado, separados por área de análise.

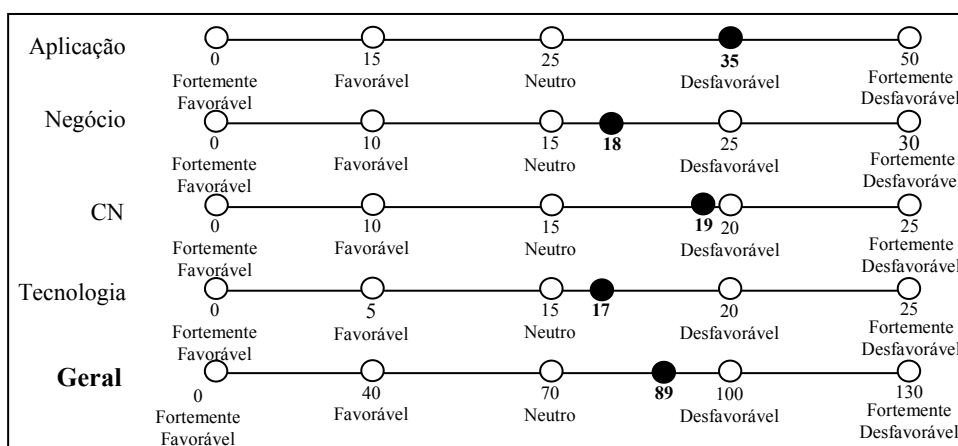


Figura 1. Análise da viabilidade de migração dos sistemas de *billing*

Após a análise da metodologia, é possível notar que o sistema de *billing* analisado não possui características nativas de nuvem e é customizado para cada cliente, os recursos não são provisionados e liberados de acordo com a demanda e grandes mudanças em sua arquitetura são necessárias para manter o desempenho e a segurança em um ambiente de computação na nuvem. Consequentemente, a evolução do sistema para a nuvem foi considerada inviável, uma vez que haveria grande impacto em sua arquitetura, e nenhum ganho significativo de expansão de clientes em potencial.

5. Conclusão

Disponibilizar o software como serviço e reduzir custos para os clientes são necessidades do mercado e justificam a popularidade do paradigma da computação na nuvem. O processo de migração de um sistema de software para a nuvem deve ser conduzido com muita cautela, sendo que antes de se optar pela migração é necessário identificar e compreender o negócio e alguns fatores técnicos que poderão viabilizar ou não a migração. Foi apresentada, neste artigo, uma metodologia que auxilia no processo de decisão da empresa quanto à evolução de um sistema de software para a nuvem.

A metodologia visa apresentar características da aplicação bem como os aspectos de negócio e tecnológicos, vinculados a modelos SaaS, que devem ser

considerados na análise da viabilidade de migração do software para a nuvem. Essa metodologia foi aplicada a um sistema de *billing*, desenvolvido pelo CPqD para empresas operadoras de telecomunicações, o que proporcionou uma análise das características tipicamente ofensoras da migração para um modelo SaaS.

Como trabalhos futuros, seria importante validar os módulos do sistema isoladamente para analisar se algum módulo pode ser evoluído para a nuvem sem grande impacto.

Referências

- Cambiucci, W. (2009) "Uma introdução ao Software + Serviços + SaaS e SOA". Disponível em: <http://msdn.microsoft.com/pt-br/library/dd875466.aspx>. Maio.
- Chauhan, A. M.; Babar, A. M. (2011) "Migrating Service-Oriented System to Cloud Computing: An Experience Report". In: 2011 IEEE 4th International Conference on Cloud Computing (CLOUD '11). IEEE Computer Society, Washington, DC, USA. **Proceedings**... 404-411.
- Hogan, M. et al. (2011) "NIST Cloud Computing Standards Roadmap". Disponível em: http://www.nist.gov/itl/cloud/upload/NIST_SP-500-291_Jul5A.pdf. Acesso em: nov. 2011.
- Hosseini, A. K. et al. (2012) The Cloud Adoption Toolkit: supporting cloud adoption decisions in the enterprise. *Software: Practice and Experience*, 42: 447465, 2011.
- Jansen, W. e Grance, T. (2011) Institute of Standards and Technology. "Guidelines on Security and Privacy in Public Cloud Computing", Special Publication 800-144. Disponível em: <http://csrc.nist.gov/publications/nistpubs/800-144/SP800-144.pdf>. Acesso em: dez. 2011.
- Mell, P.; Grace, T. (2011) "The NIST Definition of Cloud Computing: Recommendations of the National Institute of Standards and Technology". Disponível em: <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>. Acesso em: nov. 2011.
- NIST (2012) Disponível em: www.nist.gov. Acesso em: fev. 2012.
- Sousa, F. R. C.; Moreira, L. O.; Machado, J. C. (2010) "Computação em Nuvem: Conceitos, Tecnologias, Aplicações e Desafios". Universidade Federal do Ceará (UFC). Publicado no SWIB.
- Taurion, C. (2009) "Computação em Nuvem. Transformando o mundo da Tecnologia da Informação", Editado por Sergio Martins de Oliveira, Brasport.

Utilização de dispositivos móveis para o monitoramento e controle de transdutores inteligentes aplicados na automação residencial

Rogéria Oliani¹, Alexandre César R. da Silva¹, Tércio Alberto dos Santos Filho²

¹Departamento de Engenharia Elétrica – Universidade Estadual Paulista (UNESP)
Ilha Solteira – SP – Brasil

²Departamento de Ciência da Computação - Universidade Federal de Goiás (UFG)
Catalão – GO - Brasil

rooliani@gmail.com, acrsilva@dee.feis.unesp.br, tercioas@gmail.com

Abstract. *Devices such as smartphones, tablets and PCs are already part of the reality of millions of people. The use of many devices with wireless network access enables them to communicate between themselves, thus increasing the comfort and mobility of those who use them. In this paper presents a system capable of controlling and monitoring transducers, such as: light, movement sensor, electric gate motor and others devices connected in wall socket, using tablets/smartphones from wherever the user is having access to Internet.*

Resumo. *Aparelhos como smartphones, tablets e microcomputadores já fazem parte da realidade de milhões de pessoas. O uso de diversos aparelhos com acesso à rede sem fio possibilita que estes se comuniquem entre si, aumentando assim, o conforto e a mobilidade de quem os utilizam. Neste trabalho apresenta-se um sistema capaz de controlar e monitorar transdutores, como: lâmpada, sensor de movimento, motor de portão eletrônico, bem como aqueles conectados às tomadas, utilizando tablets/smartphones, de qualquer lugar em que o usuário esteja com acesso à Internet.*

1. Introdução

A tecnologia faz parte cada vez mais do dia a dia das pessoas, proporcionando-lhes conforto, segurança e mobilidade. Atualmente os dispositivos móveis (celular, PDA's, tablets, iPhones) não são mais utilizados apenas para efetuar ligações, mas também para realizar diversas tarefas, tais como: controlar a agenda de compromissos, acessar à Internet, ler e-mail, efetuar filmagens.

Semelhantes ao modo de funcionamento dos microcomputadores, os *smartphones* e *tablets* possuem sistemas operacionais, no qual realizam o interfaceamento homem máquina, facilitando ao usuário a manipulação de seus dados. Um exemplo de sistema operacional é o Android [Mass 2012].

O Android é uma plataforma *open-source* criada pela Google para o desenvolvimento de aplicações para dispositivos móveis. Inclui um sistema operacional baseado em Linux e diversas aplicações, interface gráfica, navegador web, suporte a

multimídia, GPS (*Global Positioning System*), banco de dados integrado, jogos em 3D e muito mais [Lecheta 2010].

Baseado nas características dos dispositivos móveis, este trabalho foi desenvolvido com o propósito de ampliar o conforto, a segurança e a mobilidade dentro e fora das residências, proporcionando aos seus usuários controlar a iluminação, as tomadas, os sensores de movimento e os portões automáticos de garagem de suas residências de qualquer lugar em que estejam com acesso à *Internet*, fazendo uso de *smartphones* ou *tablets* com SO (Sistema Operacional) Android. Para tanto, foi desenvolvido o software denominado SAR (Software de Automação Residencial), o qual possui dois módulos SARC (Software de Automação Residencial Cliente), desenvolvido para *tables/smarthphones*, e SARS (Software de Automação Residencial Servidor) que foi instalado em uma placa Arduino Mega 2560. Ao termino do sistema, foram realizados testes de comunicação entre o módulo do cliente e o servidor.

Na Seção 2 apresenta-se uma abordagem sobre Casas Inteligentes demonstrando algumas de suas características.

2. Casas Inteligentes

Na área de automação residencial vários trabalhos vêm sendo desenvolvidos envolvendo a rede elétrica e hidráulica de casas, equipamentos eletrônicos, bem como o mobiliário destas.

Segundo [Alves & Mota 2003], casas inteligentes são aquelas que possuem características capazes de tornar a vida, de quem nelas habita, mais simples; envolvendo aspectos, como: segurança, economia, conforto, ecologia e integração. A importância desses aspectos tem um valor distinto para cada indivíduo, em cada momento.

A gestão dos espaços internos e externos com o controle integrado de iluminação, irrigação, som ambiente, cortinas e de diversos outros equipamentos, seja local ou remoto, com o uso da *Internet*, tem conquistado cada vez mais clientes no mercado imobiliário e atendem pessoas da classe A, B e C [Loyola 2011].

Além de comodidade, a tecnologia empregada nas Casas Inteligentes pode auxiliar para a segurança de crianças e idosos, bem como ajudar pessoas com deficiência a ter maior independência [Daquino 2012].

Outras funcionalidades que podem ser aplicadas às casas inteligentes é o controle do consumo energético, possibilitando uma energia mais verde e casas mais eficientes, com o controle exato de como e quando é gasta a energia. No caso de um veículo elétrico, por exemplo, poderia calcular a parcela exata de consumo alocada ao seu abastecimento [Inovcity 2010].

3. Ferramentas utilizadas no desenvolvimento do sistema

Nesta seção apresentam-se os procedimentos utilizados para o desenvolvimento do sistema e as ferramentas utilizadas. É importante destacar que o sistema foi dividido em duas fases de desenvolvimento, sendo: programação dos dispositivos embarcados *tablets/smartphones* e a programação da placa Arduino Mega 2560 e do *Shield¹ Ethernet W5100* para o controle efetivo dos transdutores utilizando à *Internet*.

¹ *Shield* – placas de circuito contendo outros dispositivos.

Na Seção 3.1, apresentam-se as características dos *hardwares* utilizados para o desenvolvimento do módulo servidor e na Seção 3.2 apresentam-se as linguagens de programação utilizadas no desenvolvimento do software para os dispositivos embarcados, bem como para o desenvolvido para a placa Arduino.

3.1. Hardwares

Para o desenvolvimento do módulo de servidor foram utilizadas três placas: Arduino Mega 2560, *Shield Ethernet* W5100 e um *Shield* de Automação e Alarme Residencial V2.

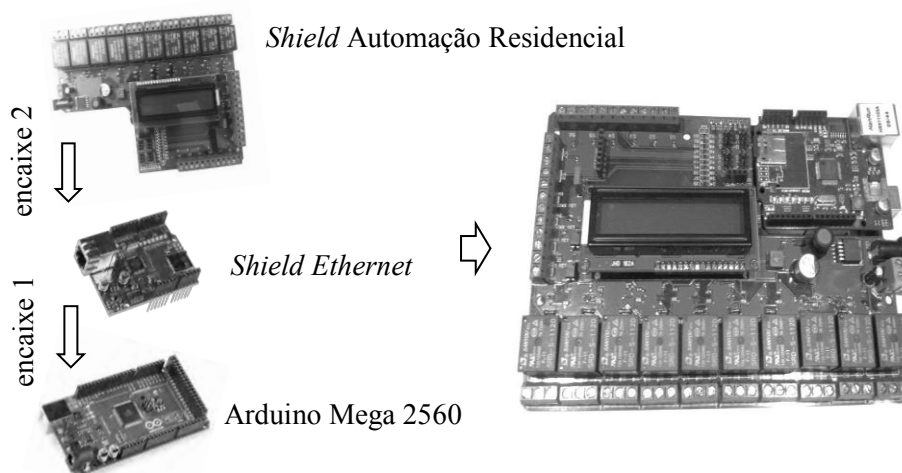
A placa Arduino Mega 2560, a qual possui um microcontrolador baseado no ATmega2560, possui 54 pinos de entradas/saídas digitais, 16 entradas analógicas, 4 UARTs (portas seriais de hardware), um oscilador de cristal de 16 MHz, uma conexão USB, uma entrada de alimentação, uma conexão ICSP (*In-Circuit Serial Programming*), um botão de reset e 256 KB de memória flash.

O *Shield Ethernet* W5100 possibilita à placa Arduino se comunicar com outros *hosts* através da *Internet*, ou até ter um servidor web embarcado, o qual viabiliza aos usuários acessarem páginas dentro do Arduino. Este *shield* fornece acesso à rede IP (*Internet Protocol*) através dos protocolos TCP (*Transmission Control Protocol*) ou UDP (*User Datagram Protocol*), bem como possui um *slot* para cartão SD que pode ser usado para armazenar arquivos [Quines 2008].

A placa de automação residencial é um circuito de condicionamento de sinal que recebe sinais elétricos de comando da placa Arduino e realiza o controle ou o monitoramento do ambiente em que esteja implementado.

Na Figura 1 apresenta-se o encaixe das placas para automação residencial; sendo, primeiramente, conectado o *Shield Ethernet* à placa Arduino Mega 2560 e, posteriormente, estas ao *Shield* de Automação e Alarme Residencial V2.

Figura 1 - Encaixe das placas para automação residencial.



Fonte: Elaborada pelo autor.

Para o desenvolvimento deste trabalho, bem como para efetuar os devidos testes, além da placa Arduino Mega 2560, *Shield Ethernet* W5100 e *Shield* de Automação e Alarme Residencial V2 foram utilizados os seguintes equipamentos: um roteador; um modem; duas lâmpadas; uma tomada; um sensor de movimento; um controle de portão

eletrônico; um *tablet* com Sistema Operacional Android; um *smartphone* com Sistema Operacional Android.

3.2. Linguagens de programação

Foram utilizadas no desenvolvimento deste trabalho as linguagens de programação Arduino e Java. A linguagem Arduino foi utilizada na programação da placa Arduino e a linguagem Java, em conjunto com o Android SDK (*Software Development Kit*), no desenvolvimento do sistema manipulado pelo usuário no *tablet/smartphone*.

A linguagem Arduino é baseada em linguagem C, e utiliza um software livre, Arduino v 1.0.1, para a escrita dos códigos do sistema criado para a placa Arduino, bem como para efetuar o *upload* na placa deste. A linguagem C é uma linguagem genérica que, apesar de ser de alto nível, compartilha recursos tanto de alto nível, quanto de baixo, pois permite acesso e programação direta de microprocessadores [Dornelles 1997].

O Android SDK possui uma biblioteca API e uma ferramenta para desenvolver, testar e debugar aplicativos para Androids utilizando a linguagem Java. Uma das grandes vantagens é a diversidade de aparelhos encontrados no mercado utilizando o sistema operacional Android. Atualmente são várias as empresas que desenvolvem referidos aparelhos, como, por exemplo: Motorola, LG, Samsung, HTC e outros. Outra característica relevante é que o mesmo possui o código aberto, e permite uma programação personalizada, sendo possível customizar os aparelhos baseados em sua característica.

4. Desenvolvimento do Sistema de Automação Residencial (SAR)

O Sistema de Automação Residencial desenvolvido foi denominado de SAR foi dividido em dois módulos:

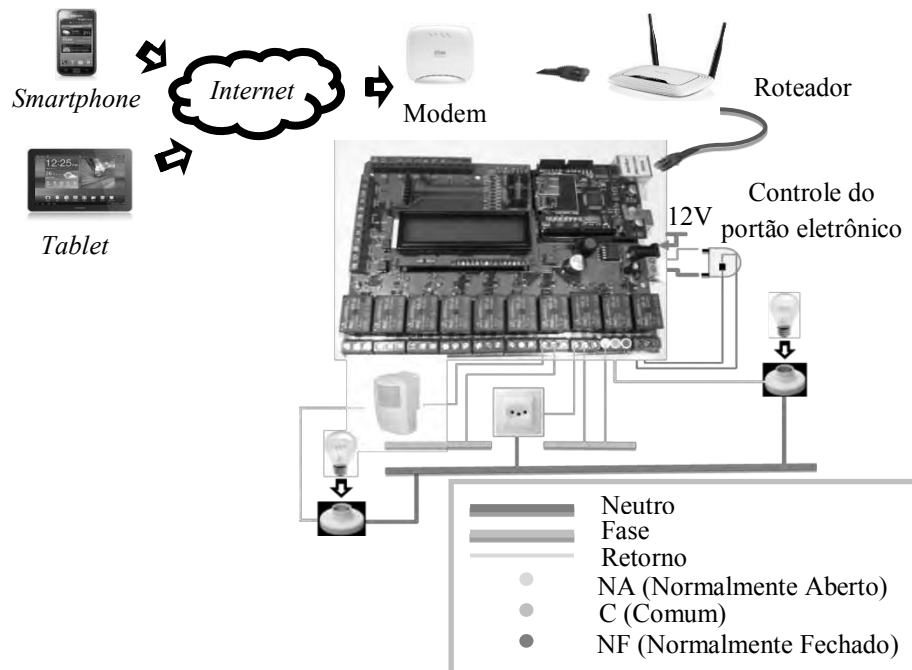
1. SARS (Software de Automação Residencial Servidor) - desenvolvido para a placa Arduino;
2. SARC (Software de Automação Residencial Cliente) - desenvolvido para *tablet/smartphone* com SO Android.

Conforme apresentado na Figura 2, o SARC enviará a requisição solicitada pelo usuário através da *Internet*. Por intermédio do modem, do roteador e do *Shield Ethernet*, esta requisição chegará à placa Arduino na qual o SARS a interpretará e ativará/desativará o relé localizado no *Shield* de Automação Residencial, conforme solicitado.

Na placa de Automação Residencial estão sendo utilizados quatro relés. Na Figura 2, visualizando da direita para a esquerda, no primeiro relé está conectado o controle de portão eletrônico, no segundo a lâmpada, no terceiro a tomada e no quarto o sensor de movimento.

Os relés funcionam como interruptores que permitem ou não que a tensão da rede elétrica da residência chegue até os transdutores, permitindo que se acenda/apague a lâmpada, ligue/desligue o sensor de movimento, a tomada e o controle do portão eletrônico.

Figura 2 - Esquemático do sistema implementado.



Fonte: Elaborada pelo autor.

Para a requisição realizada pelo usuário através do *tablet/smartphone* trafegue pela *Internet* e chegue ao seu destino, é necessário que este seja identificado através de seu IP. O IP identifica um dispositivo tanto na rede local, quanto na rede mundial de computadores. O *Shield Ethernet* utilizado neste trabalho é identificado na rede local por um IP, o qual foi determinado no software e programado na placa Arduino. Assim, quando receber um dado da rede externa o roteador consegue localizá-lo na rede local, realizando um NAT (*Name Address Translate*), pois, a *Internet* e a rede local são redes distintas e possuem IPs dedicado a sua rede.

Neste trabalho os testes foram realizados através de uma conexão utilizando IP dinâmico. Para tanto, foi necessário que o destino das requisições realizadas através do *tablet/smartphone*, por IP, fosse realizada através do uso de DNS (*Domain Name System*).

Os Serviços Dinâmicos DNS - como o do No-IP, o qual foi utilizado neste trabalho, possibilitam que um dispositivo seja localizado na *Internet*, mesmo com a mudança de seu IP (IP Dinâmico), pois estes sistemas, através de seus aplicativos, informam de tempos em tempos o IP atual, efetuando automaticamente a atualização do redirecionamento.

4.1. SARS - Software Desenvolvido para a placa Arduino (servidor)

Como sua própria denominação específica, o SARS atua como um Servidor Web recebendo as requisições que são enviadas pelo cliente (SARC), executando processos, conforme as requisições recebidas, e enviando respostas aos clientes, com ou sem dados.

O trabalho foi desenvolvido para efetuar o controle de quatro relés, os quais serão ativados ou desativados conforme as requisições enviadas pelo software cliente,

bem como informar o status (ativado/desativado) de três destes relés (relés que controlam o sensor, lâmpada e tomada). O relé que ativa o controle remoto do portão eletrônico não fornece seu status, em virtude do mesmo ser desativado 200 ms (milissegundos) após a sua ativação.

Por questão de segurança, as requisições feitas pelos clientes são autenticadas (usuário e senha) e enviadas utilizando-se o método POST através do uso do protocolo de transferência HTTP (*HyperText Transfer Protocol*).

4.2. SARC - Software desenvolvido para *Tablet/Smartphone* (cliente)

O SARC foi desenvolvido para o Sistema Operacional Android (*Tablet/Smartphone*), o qual é utilizado pelo usuário para efetuar o controle dos transdutores conectados a placa de Automação Residencial, e atua como um cliente, em uma relação cliente/servidor, com o software desenvolvido para a placa Arduino (SARS).

Na Figura 3 apresenta-se a *Interface* do SARC, na qual é possível identificar sete botões: Tomada – ativa ou desativa a tomada; Lâmpada – acende ou apaga a lâmpada; Sensor – liga ou desliga o sensor; Garagem – aciona o controle do portão eletrônico da garagem por 20 ms; *Internet* – ativa ou desativa o uso da *Internet* pelo sistema; Atualizar – habilita ou desabilita a atualização do status do SARS a cada 1 segundo; Sair – encerra a aplicação.

Figura 3 - Interface do SARC.



Fonte: Elaborada pelo autor.

Ao lado dos botões “*Internet*” e “*Atualizar*” são exibidas as seguintes informações:

- Tempo de Resposta – tempo decorrido entre o envio da requisição e o retorno de sua resposta, em segundos;
- Tempo Médio – tempo médio de resposta;
- Atualizações – número de atualizações realizadas.

5. Resultados

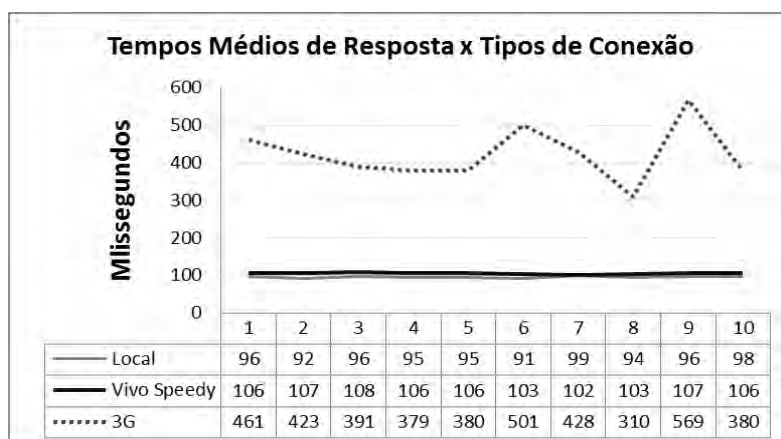
Com o objetivo de verificar a viabilidade do Sistema de Automação Residencial (SAR) no envio e recebimento de dados, foram efetuados testes utilizando rede local, *Internet* Banda Larga Vivo Speedy (plano de 1 Mb/s) e *Internet* Móvel 3G Vivo.

O Tempo Médio de Resposta (TMR) mínimo obtido foi 91 ms (rede local) e o máximo 569 ms (*Internet* Móvel 3G Vivo). Apesar da expressiva diferença entre o mínimo e o máximo alcançado nos diferentes tipos de conexões utilizados, o sistema apresentou um tempo de resposta viável. No tempo de resposta mais lento, o usuário aguardou 0,56 segundo para obter uma resposta, o que representa mais da metade do tempo que o sistema leva para efetuar a atualização automática, que foi determinada para ocorrer a cada segundo.

Para a realização das análises foram realizadas dez medições do Tempo Médio de Resposta, obtido através de cem atualizações cada, totalizando mil atualizações (repostas obtidas), para cada tipo de conexão (Local, *Internet* Vivo Speedy e 3G Vivo). O *Shield Ethernet* conectado a placa Arduino permaneceu em todos os testes comunicando-se com a *Internet* por meio da conexão Vivo Speedy, exceto quando da realização dos testes utilizando-se conexão local. Independente do tipo de conexão, as requisições foram realizadas e tiveram um retorno de cem por cento dos dados.

No Gráfico 1 apresenta-se os resultados dos testes realizados para obtenção do TMR do SAR, no qual se pode observar a oscilação tida com a conexão 3G Vivo, bem como os resultados obtidos com a conexão Local e Speedy, as quais apresentaram pouquíssimas variações, com um TMR mínimo de 91 e 102 ms, e um máximo de 99 e 108 ms, respectivamente.

Gráfico 1 - Resultado dos testes realizados para obtenção do TMR do SAR.



Fonte: Elaborada pelo autor.

6. Conclusão e Trabalhos Futuros

O trabalho desenvolvido resultou em um sistema denominado SAR, no qual foram efetuados testes que demonstraram a sua viabilidade no tocante ao envio e recebimento de dados.

O controle e monitoramento de sensor, lâmpada, tomada e controle de portão eletrônico apresentaram-se eficazes, ao ponto que todas as requisições recebidas pelo SARS foram executadas de maneira satisfatória.

Assim, com base nos testes realizados, o SAR demonstrou ser um sistema viável, e que pode ser expandido com a agregação de outros transdutores e funções, com foco nas pessoas e no meio ambiente, propiciando mais conforto a aquelas, e respeito a este, com a utilização responsável da energia e das águas.

Como possíveis trabalhos futuros pode-se apontar a implementação de um Banco de Dados utilizando uma estrutura de dados JSON (*JavaScript Object Notation*), a ser armazenado em um cartão micro SD (*Shield Ethernet*), contendo os dados dos usuários, tipos de requisições, data/hora dos acessos a fim de se manter um histórico destes, bem como os tipos de usuários, controlando, assim, o nível de acesso ao sistema.

Agradecimentos

Os autores agradecem ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (Processo: 307255/2009-3) e a Universidade de Jales – UNIJALES.

Referências

- Alves, José Augusto, e José Mota. *Casas Inteligentes*. 2003. <http://www.centroatl.pt/titulos/solucoes/casasinteligentes.php3> (acesso em 11 de 11 de 2012).
- Daquino, Fernando. *Casas inteligentes serão realidade em 10 anos, afirmam cientistas*. 31 de março de 2012. <http://www.tecmundo.com.br/casas/21501-casas-inteligentes-serao-realidade-em-10-anos-afirmam-cientistas.htm> (acesso em 11 de novembro de 2012).
- Dornelles, Adalberto A. “Fundamentos de Linguagem C.” *Departamento de Engenharia de Computação e Automação - UFRN*. 1997. http://www.dca.ufrn.br/~xamd/dca0800/apostila_C.pdf (acesso em 09 de novembro de 2012).
- Inovcity. *Casas mais Inteligentes*. 2010. <http://www.inovcity.pt/pt/Pages/casas-inteligentes.aspx> (acesso em 11 de novembro de 2012).
- Lecheta, Ricardo R. *Google Android Aprenda a criar aplicações para dispositivos móveis com o Android SDK*. São Paulo: Novatec, 2010.
- Loyola, Darshany. *'Casas inteligentes' atraem cada vez mais clientes no mercado capixaba*. 16 de setembro de 2011. <http://g1.globo.com/espírito-santo/mercado-imobiliario/noticia/2011/09/casas-inteligentes-atraem-cada-vez-mais-clientes-no-mercado-capixaba.html> (acesso em 11 de novembro de 2012).
- Mass, FRAMINGHAM. *Android Marks Fourth Anniversary Since Launch with 75.0% Market Share in Third Quarter, According to IDC*. 1 de novembro de 2012. <https://www.idc.com/getdoc.jsp?containerId=prUS23771812> (acesso em 2012 de novembro de 15).
- Quines, Rafael. *EthernetShield Oficial do Arduino*. 8 de novembro de 2008. <http://www.arduino.net/blog/tag/ethernet-shield-arduino-internet-servidor-web/> (acesso em 10 de outubro de 2012).

Desempenho do Uso de Visões nas Consultas em Diferentes Servidores de Banco de Dados Relacionais

Walter José da Silva¹, Hiago de Assis Silva¹, Flávio Ferreira Borges²

¹Universidade Federal de Goiás – Campus Jataí
BR 364, Km 193, nº 3800, CEP: 75801-615 – Jataí – GO – Brasil

²Departamento de Ciências da Computação
Universidade Federal de Goiás – Campus Jataí – Jataí, GO – Brasil
{sillva.walter, hiagoassis, flavio.ufg.jatai}@gmail.com

Abstract. *This paper presents a research on the performance of queries using view in the database, to identify whether the use of this feature can bring benefits to alleviate the overload of requests for database servers, reducing the response time. And also which will show the performance gain from using this feature for different database management systems in order to know if there is a performance difference and see which of the servers tested support better the feature of view.*

Resumo. *Este artigo apresenta um trabalho de investigação do desempenho do uso de visão em consultas ao banco de dados, visando identificar se o uso desse recurso pode trazer benefícios para amenizar a sobrecarga de requisições em servidores de banco de dados, diminuindo o tempo de resposta. E também mostrar qual será o ganho de desempenho do uso desse recurso para diferentes sistemas gerenciadores de banco de dados, a fim de saber se há uma diferença de desempenho e verificar qual dos servidores testados suporta melhor o recurso de visão.*

1. Introdução

A evolução das organizações e o aumento no volume de dados manipulados ocasionaram o surgimento de uma série de sistemas específicos para tratar as várias necessidades de um banco de dados. Dentre essas necessidades destacam-se o controle do acesso aos dados, o aperfeiçoamento do desempenho e a garantia da integridade. Para facilitar esta administração surgiu o Sistema Gerenciador de Banco de Dados (SGBD) [Silberchatz *et al.* 2006].

O SGBD é considerado como uma coleção de programas que permite aos usuários a criação e manutenção de um banco de dados, portanto, é um sistema de *software* de propósito geral que facilita os processos de definição, construção, manipulação e compartilhamento de um conjunto de dados entre vários usuários e aplicações [Elmasri *et al.* 2005].

O meio de acesso às informações mantidas por esse SGBD é conhecida por *query*, que é simplesmente uma forma de requisitar informações ou puramente dados ao banco. Essa forma de requisitar informações é possível graças a um padrão criado no

início da década de 70, chamado de SQL (*Structured Query Language*) [Date 2003]. A SQL é uma linguagem para lidar com bancos de dados relacionais, e que abrange instruções para definição de dados, consultas e atualizações, dentre outras funcionalidades [Date 2003], [Elmasri *et al.* 2005], [Ramakrishnan *et al.* 2008].

Porém, essas requisições consomem certo tempo de processamento, e em sistemas de grande porte, onde o número de requisições simultâneas é elevado, o servidor de banco de dados pode encontrar dificuldades em enviar as repostas aos processos chamadores (clientes) em um tempo hábil. E para amenizar a sobrecarga dos servidores no processamento dessas requisições sugere-se o uso de visão em consultas ao banco de dados.

Em banco de dados, visão é um recurso avançado onde algumas de suas funções são fornecer um meio de abreviação ou de macro, permitir que vários conjuntos de dados (tabelas) sejam vistos por usuários em um novo conjunto, fornecer segurança automática para dados ocultos e independência de dados lógica [Date 2003]. E que segundo Date (2003), Elmasri (2005) e Halevy (2001), trazem um ganho significativo no tempo de resposta da consulta.

Contudo, o objetivo deste trabalho é verificar o desempenho em tempo de resposta das consultas com o uso do recurso de visão e compará-lo ao desempenho de consultas sem o uso desse recurso, para comprovar de fato a sua eficiência. E também comparar o desempenho de visões em diferentes SGBDs, a fim de mostrar se há uma diferença relevante no tempo de resposta entre eles.

Para avaliar o desempenho das consultas com o recurso de visão de SGBDs é usado à técnica de *benchmark*. E conforme Ciferri (1995) essa técnica é baseada em um modelo de análise, composta por um conjunto de testes já previamente definidos e que poderá ser aplicado a vários sistemas distintos, que possuam funcionalidades semelhantes. O objetivo é manter toda a avaliação de desempenho dentro das mesmas características. O JMeter é ferramenta usada neste trabalho, e possui essas técnicas mencionadas.

O Apache JMeter é uma aplicação *desktop*, projetada para testar e medir o comportamento do desempenho funcional de aplicações cliente/servidor, como aplicações *web*, banco de dados ou aplicativos de FTP (*File Transfer Protocol*). É uma das ferramentas (*open-source*) amplamente utilizada em testes de sistema, oferecendo à aplicação de testes em rede livremente distribuída. É puramente baseado em Java e é altamente extensível através de uma API (*Application Programming Interface*) fornecida pela Apache [Halili 2008], [Santos *et al.* 2008].

Todavia, autores como Pires *et al.* (2006), Silva (2006) e Aquino (2007) mostram a importância de avaliar recursos ou certas funcionalidades em sistemas computacionais, que nesse caso específico são servidores de banco de dados. O motivo é verificar se há diferenças de desempenho em funcionalidades de sistemas, visando encontrar a ferramenta certa para uma aplicação específica.

Enfim, tem-se visto que essa investigação será importante para obtenção de métodos que diminuirão um problema tão comum em dias atuais, onde a geração de informação armazenada cresce demasiadamente, e a busca por essas informações chega a ser frenética. Técnicas para amenizar a sobrecarga de servidores de dados são

importantes, e o recurso de visão pode ser mais uma dentre outras várias técnicas existentes a realizar essa tarefa.

2. Trabalhos Relacionados

O objetivo que se tem com esses trabalhos é apresentar as discussões e problemas relacionados ao uso de visão em consultas ao banco de dados, e encontrar formas e métodos de se fazer uma avaliação de desempenho entre *softwares* distintos. De forma que essa avaliação seja realizada em um alto nível de imparcialidade, buscando o embasamento necessário para mostrar a relevância de se fazer essa análise.

Começando com o trabalho de Mitra (2001) onde é apresentado um algoritmo que permite que um sistema que manipula uma quantidade muito grande de informações possa reescrever consultas de forma eficiente utilizando visão. Com isso, reduzir o tempo de resposta aos usuários, melhorar a escalabilidade e tornar o sistema capaz de lidar com um grande número de fontes de informação¹.

Nos trabalhos de Halevy (2001) e Felea *et al.* (2007) é mostrado a importância de usar visão em consultas ao banco de dados. É dito que visão é uma funcionalidade que ajuda a resolver uma série de problemas de gerenciamento de dados. A otimização da consulta, a manutenção da independência física dos dados, a integração de dados e o projeto de *design* de armazenamento de dados são os problemas citados por Halevy (2001), que podem ser resolvidos por meio do uso de visão.

Halevy (2001) ainda ressalta que as muitas aplicações da visão usadas para responder consultas estimularam uma onda de pesquisa, que foram desde os fundamentos teóricos até o projeto do algoritmo e a sua implementação em vários sistemas comerciais. E Felea *et al.* (2007) reuni pontos de vista de vários autores para falar sobre os problemas que foram solucionados com o uso de visão.

O problema em servidores de banco de dados sobrecarregados é um problema abordado por Kyoung-Don *et al.* (2007), ele fala principalmente em aplicações de tempo real onde o tempo de resposta para uma requisição é um fator crítico. Com base em estudos apresentados por autores já citados, o uso de visão em consultas ao chamado RTDB (*Real-Time Database*), poderia ser uma solução para amenizar a sobrecarga, porém não mencionada pelo autor, que aborda um outro tipo de tecnologia chamada de *Chronos*. O trabalho de Kyoung-Don *et al.* (2007) foi encarado aqui simplesmente como um exemplo de que existe o problema de sobrecarga em servidores de banco de dados e é preciso realizar trabalhos que ajudem a diminuir esse problema.

Com relação à avaliação de desempenho de sistemas, o trabalho de Pires *et al.* (2006) tem a proposta de apresentar um estudo comparativo de desempenho entre os servidores de banco de dados MySQL e PostgreSQL, com a utilização da ferramenta *Open Source Database Benchmark* (OSDB), em uma plataforma GNU/Linux. O trabalho consiste em analisar os SGBDs sob as métricas geradas pelo OSDB e fomentar possíveis melhorias em suas funcionalidades.

¹ Informações de pesquisas recorrentes de usuários [Mitra 2001].

De outra forma, o trabalho de Silva (2006) tem o objetivo de avaliar o desempenho dos principais métodos de armazenamento de objetos multimídia, com uso de *benchmark*. E usa como estudo de caso para avaliação o SGBD PostgreSQL.

Devemos também ressaltar o trabalho de Aquino (2007), que apresenta o sistema VEPersonal (*Personalized Virtual Environment*), uma infraestrutura para gerenciamento de componentes de realidade virtual, o qual permite adaptar ambientes virtuais em função do nível de conhecimento ou das necessidades do usuário. E teve o objetivo de propor, construir e validar uma infraestrutura para fazer o gerenciamento de componentes de ambientes virtuais adaptativos.

No entanto, um dos problemas principais do trabalho é a construção de mundos virtuais utilizando objetos 3D armazenados em um SGBD. Contudo, o armazenamento e a recuperação de objetos em um SGBD podem ser realizados por linguagens de consulta com suporte a XML. E partir disso, surge à necessidade de comparar SGBDs com XML nativo com SGBDs com suporte a XML [Aquino 2007].

Enfim, é notável a relevância da investigação aqui proposta, já que visão é um recurso que possui uma capacidade de diminuir um problema muito comum em servidores de grande porte, que é a sobrecarga. E conforme mostrado em trabalhos anteriores, existe uma necessidade de analisar e comparar tecnologias diferentes sob certos aspectos funcionais. E que essas análises sejam feitas de forma automatizada e metodológica, para que haja uma relevância e coerência na decisão de escolha do *software* apropriado para cada caso.

3. Metodologia

A metodologia usada na avaliação dos SGBDs foi a GQM (*Goal - Question - Metric*). No qual a ideia básica é derivar métricas de *software* a partir de questões que queremos responder e objetivos que queremos alcançar [Guarizzo 2008]. Com o intuito de mostrar quais das métricas fornecidas pela ferramenta JMeter vão ser utilizadas neste trabalho.

A seguir são especificados os pontos GQM dessa metodologia:

- **Objetivo (*Goal*):** o objetivo desse trabalho é descobrir se as consultas com recurso de visão podem trazer um ganho de desempenho em tempo de resposta para servidores de banco de dados. E também investigar se consultas com visão podem ter desempenhos diferentes em SGBDs distintos.
- **Questões (*Question*):** Consultas com visão são melhores que consultas sem esse recurso? Consultas com visão podem trazer um ganho significativo para servidores de banco de dados sobrecarregados?
- **Métricas (*Metric*):** Com base nos objetivos e nas questões levantadas, percebe-se que as métricas **média** e **série 90%**², fornecidas pelo JMeter, são suficientes para gerar o conhecimento necessário para atingir os objetivos. Essas medidas serão usadas para compor as fórmulas abaixo.

² Essa métrica mostra que 90% das amostras atingiram um determinado tempo.

Média Harmônica: serve para achar o *throughput* por unidade de tempo de cada requisição. Ela é definida pela fórmula mostrada logo a seguir, onde n é o número de requisições, e t é o tempo de resposta para cada requisição [Silberchatz *et al.* 2006].

$$Throughput = n / ((1 / t_1) + (1 / t_2) + \dots + (1 / t_n))$$

Speedup: ajuda a mostrar o ganho da diferença de desempenho entre as consultas, ou seja, quanto maior o *Speedup*, maior será o ganho de desempenho. É dado pela fórmula a seguir:

$$Speedup = oldRequestTime / newRequestTime$$

4.1. Descrição dos Testes

Todos os testes foram realizados em um notebook DELL® Inspiron 1545, com as seguintes configurações: processador Intel® Pentium Dual-Core CPU T4200 @ 2.00GHz, 2000Mhz, 2 núcleos, 2 processadores lógicos, com 512KB de *cache*, RAM de 3GB e disco SATA de 150 GB e 5300 rpm.

A etapa inicial compreende na instalação dos servidores de banco de dados MySQL 5.5.29 e o PostgreSQL 8.3. Os dois servidores foram instalados no sistema operacional Windows, e como o sistema operacional não influencia no resultado dos testes, esse foi escolhido meramente por proporcionar comodidade ao pesquisador.

A escolha dos SGBDs se deve ao fato de serem os mais utilizados pelos usuários, o que é refletido nas várias pesquisas feitas sobre eles, mostradas inclusive nos trabalhos relacionados. E também por termos permissão para realizar testes e divulgá-los, o que não seria possível em servidores de banco de dados proprietários, em que por questões de *marketing*, não permitem a divulgação dos resultados [Pires *et al.* 2006].

Para realizar os testes foi usada uma base de dados de gerenciamento escolar, já igualmente predefinida para os dois SGBDs, que contém 103 tabelas e 97 relacionamentos, e com um tamanho de aproximadamente 135MB. Em seguida foram construídas 4 consultas para os dois servidores, com 5, 10, 15 e 20 relacionamentos cada.

4.2. Carga de Trabalho

Pretende-se com as cargas de trabalho, esforçar ao máximo o servidor, tentando simular uma condição de uso real e observar o seu comportamento. Elas foram impostas aos dois SGBDs testados em 4 fases, como seguem abaixo:

- 1º Fase: Foram executadas as 4 consultas sem o recurso de visão (CTs) com 4, 8, 16, 32, 64, 128 usuários simultâneos de cada vez, no MySQL;
- 2º Fase: No PostgreSQL foram executadas as 4 (CTs) com a mesma quantidade de usuários simultâneos por vez;
- 3º Fase: Depois, foram executadas no MySQL as 4 consultas com o recurso de visão (CVs) também com 4, 8, 16, 32, 64, 128 usuários simultâneos;
- 4º Fase: E por fim, executamos no PostgreSQL as 4 (CVs), também com a mesma quantidade de usuários simultâneos descrita acima.

É importante enfatizar que tanto as (CTs) e as (CVs) construídas nos dois SGBDs, não usam o recurso de índices. Os índices são usados pelos servidores de banco de dados para obter um acesso aleatório rápido às tuplas do banco, onde cada índice está associado a uma determinada chave de busca [Silberchatz *et al.* 2006].

4. Resultados Obtidos

Para realizar os testes, alguns cuidados foram tomados para que não fossem considerados certos dados que não seguem um padrão em relação aos demais. Para isso foi usado a métrica série 90% (fornecida pelo JMeter), e comparou-se com a média dos tempos de resposta das requisições. Só foram validadas as médias que ficaram entre um intervalo de 10% (para mais ou para menos), do resultado gerado pela série 90%. Então se afirma que no mínimo 80% dos tempos de resposta estavam bem próximos à média.

Após isso, já com todos os resultados em mãos, foi utilizado a Média Harmônica para encontrar o *throughput* para cada tempo de resposta retornado. E para saber qual servidor de banco de dados teve um maior desempenho nas consultas com visão é usada a fórmula de *Speedup*. Com ela é possível quantificar quanto de ganho se obteve em relação aos dois tipos de consulta.

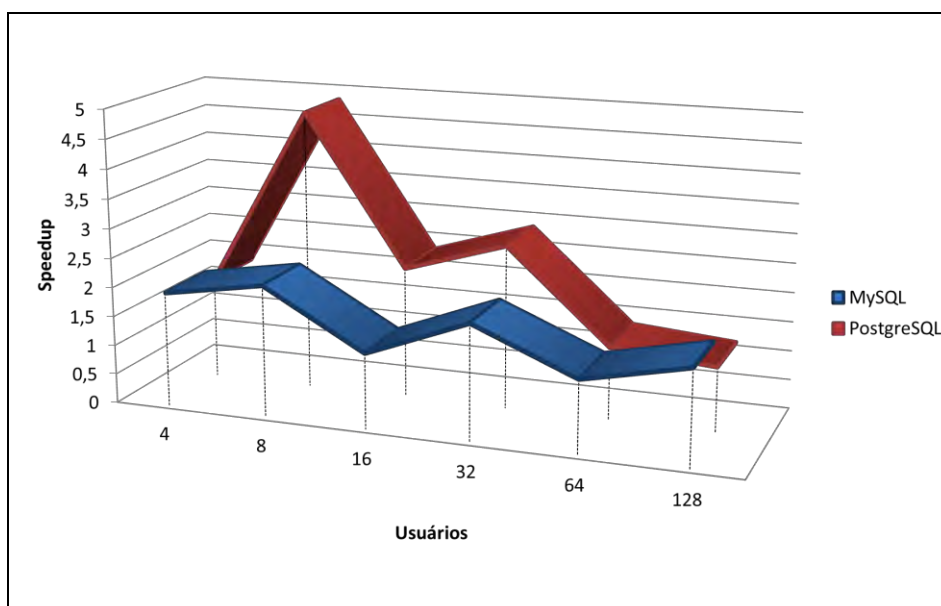


Figura 1. Comparação de *Speedup* das consultas nos SGBDs.

Na Figura 1 é mostrado um comparativo do *Speedup* entre o MySQL e o PostgreSQL, onde foi feita uma média do ganho das consultas com os diferentes relacionamentos. É possível notar que o PostgreSQL se mostrou com um resultado melhor que seu concorrente.

Por fim, é mostrada na Figura 2 a diferença total de ganho entre os dois SGBDs. O MySQL obteve um ganho de aproximadamente 70%, e o PostgreSQL obteve um ganho de 130%. Portanto, conclui-se que o PostgreSQL suporta melhor o recurso de visão, e tem uma diferença significativa de desempenho comparado com o MySQL.

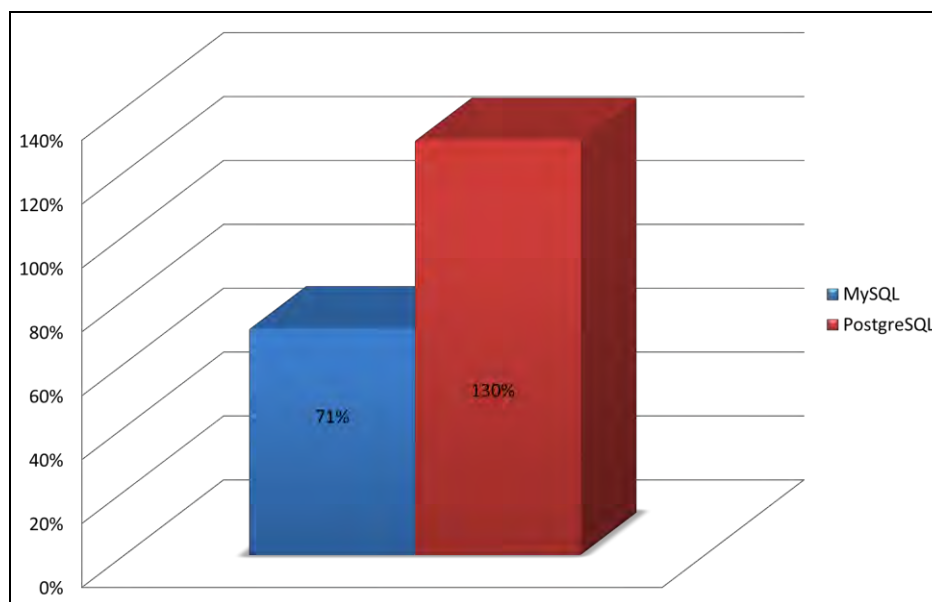


Figura 2. Desempenho das visões nos SGBDs.

5. Conclusão e Trabalhos Futuros

A sobrecarga de consultas em banco de dados é um problema que tende a se agravar com o passar do tempo. A enorme quantidade de informações requisitadas por dia pode ultrapassar a casa dos trilhões³. Desta forma, é necessário o uso de técnicas que minimizem esse intervalo de tempo, entre requisição e resposta, e alivie o meio de comunicação utilizado.

Baseado nisso, tem-se visto que o recurso de visão conseguiu retornar valores expressivos para redução de tempo de resposta, e podendo assim, contribuir com uma diminuição significativa da sobrecarga de servidores de dados. Além disso, foi possível observar que existem diferenças de desempenho de funcionalidades específicas entre SGBDs distintos, e a investigação será de grande ajuda para quem está em busca de uma ferramenta para um tipo de aplicação que se preocupa na maioria do tempo a responder requisições.

Como trabalhos futuros, é pretendido realizar esses mesmos testes sob consultas em diferentes SGBDs, mas agora usando visões materializadas, onde é guardado o resultado da consulta em vez do resultado do gerador de código, no qual pode-se observar se elas possuem diferenças de desempenho das consultas com recurso de visão virtual, a qual foi abordada neste trabalho. E verificar também como seria o seu comportamento em meio às atualizações em tempo real.

Também é desejável, aprofundar mais nas atividades exercidas pelo SGBD para responder a uma requisição, principalmente investigar o funcionamento do otimizador da consulta, e ver qual é a dificuldade de se chegar a um caminho ótimo em um tempo razoável, com o objetivo de verificar qual seria o motivo de diferentes SGBDs terem um tempo de resposta variável para consultas idênticas.

³ <http://www.realtimestatistics.org/>

References

- Aquino, Marcus S. (2007) “VEPersonal - Uma Infra-Estrutura para Geração e Manutenção de Ambientes Virtuais Adaptativos”. Tese de Doutorado, Universidade Federal de Pernambuco, Recife, PE.
- Ciferri, Ricardo R. (1995) “Um Benchmark Voltado à Análise de Desempenho de Sistemas de Informações Geográficas”. Dissertação de Mestrado, Universidade Estadual de Campinas, Campinas, SP.
- Date, C. J. (2003), Introdução a Sistemas de Banco de Dados, Editora Elsevier, 4ª edição.
- Elmasri, Ramez; Navathe, Shamkant B. (2005), Sistemas de Banco de Dados, Editora Pearson Addison Wesley, 4ª edição.
- Felea, Victor; Balta, Marian. (2007) “On Equivalence Of Queries Using Views”, In: International Symposium on Symbolic and Numeric Algorithms for Scientific Computing, IEEE Conference Publications.
- Guarizzo, Karina. (2008) “Métricas de Software”. Monografia de Graduação, Faculdade de Jaguariúna, Jaguariúna, SP.
- Halevy, A. Y. (2001) “Answering Queries Using Views: A Survey”, In: Department of Computer Science and Engineering, University of Washington.
- Halili, Emily H. (2008), Apache JMeter: A Practical Beginner's Guide to Automated Testing and performance measurement for your websites. Editora Packt Publishing, 1ª edição.
- Kyoung-Don, K.; Sin, P. H.; Jisu, O.; Son, S. H. (2007) “A Real-Time Database Testbed and Performance Evaluation”, In: 3th IEEE International Conference, Embedded and Real-Time Computing Systems and Applications.
- Mitra, Prasenjit. (2001) “An Algorithm for Answering Queries Efficiently Using Views”, In: 12th Australasian Database Conference, IEEE Conference Publications.
- Pires, Carlos E. S.; Nascimento, Rilson O.; Salgado, Ana C. (2006) “Comparativo de Desempenho entre Bancos de Dados de Código Aberto”, In: Escola Regional de Banco de Dados, Anais da ERBD06, Porto Alegre.
- Ramakrishnan, Raghu; Gehrke, Johannes. (2008), Sistemas de Gerenciamento de Banco de Dados. Editora McGraw-Hill, 3ª edição.
- Santos, Ismaley S.; Neto, Pedro A. S. (2008) “Automação de Testes de Desempenho e Estresse com o JMeter”, In: II ERCEMAPI - Escola Regional de Computação, São Luiz.
- Silberchatz, Abraham; Korth, Henry F.; Sudarshan, S. (2006), Sistemas de Banco de Dados. Editora Elsevier, 5ª edição.
- Silva, Ricardo C. (2006) “Benchmark em Banco de Dados Multimídia: Análise de Desempenho em Recuperação de Objetos Multimídia”. Dissertação de Mestrado, Universidade Federal do Paraná, Curitiba.

Ambiente Virtual de de Cenário Urbano para Simulação Multiagentes - Experiência do Projeto ADGEPA

Mirelle Cristina Lima Silva¹, Erika Moraes Martins Coelho¹, Vinícius Sebba Patto¹

¹Instituto de Informática – Universidade Federal de Goiás (UFG) – Bloco IMF-1
Campus Samambaia – CEP. 74690-815 – Goiânia – GO – Brasil

mirellycristina.si@gmail.com, {erikamoraes, viniciussebba}@inf.ufg.br

Abstract. *One of the ADGEPA project artifacts is a responsive web application in which one can make complaints about socioenvironmental concerns and can store these complaints in a data base. From these data, we look for useful knowledge, using data mining techniques. Since experiments to gather data take too long, we propose the use of multiagents systems to simulate our application users. In this paper, we present the ADGEPA project, its simulation model, and our virtual environment proposal for this simulation.*

Resumo. *Um dos artefatos do projeto ADGEPA é uma aplicação web responsiva em que usuários podem fazer denúncias sobre problemas socioambientais e as armazenar em uma base de dados. A partir dos dados, nós buscamos conhecimento útil, por meio de técnicas de mineração de dados. Haja vista que experimentos para captar dados tomam muito tempo, propomos o uso de sistemas multiagentes para simular usuários da nossa aplicação. Neste artigo, nós apresentamos o projeto ADGEPA, o seu modelo de simulação e a nossa proposta de um ambiente virtual para a nossa simulação.*

1. Introdução

À medida que cidades crescem, aumenta-se a necessidade de recursos naturais e serviços para manter suas populações. O abastecimento de recursos básicos e o fornecimento de serviços (e.g. coleta de esgoto) se tornam mais difíceis. Consequentemente, problemas sociais e ambientais são evidenciados.

Faz certo tempo que pesquisadores propõem formas de identificar e classificar áreas de risco socioambiental, apoiar processos de tomada de decisão, permitir análise conjunta de dados socioambientais, etc. Podemos citar alguns exemplos: (i) em [Alves 2006], o objetivo da pesquisa foi de operacionalizar a categoria vulnerabilidade social pelo mapeamento de pontos críticos de fragilidades socioambientais; (ii) [Barcellos et al. 2005] propuseram, por meio de técnicas de geoprocessamento, localizar casos de dengue e identificar fatores socioambientais que caracterizam locais com casos frequentes; e (iii) [Poch et al. 2004], no qual autores descrevem a construção de sistemas de apoio a decisão voltados para escolha de sistemas de tratamento de esgoto.

Trabalhos recentes continuaram a propor avaliações de riscos socioambientais, formas de apoiar ações para minimizar problemas socioambientais, etc. Muitos são pautados apenas em métodos científicos e matemáticos para atingir seus objetivos, e.g.: [Katukiza et al. 2010] e [Malekpour et al. 2013]. Outros consideram também a participação pública em processos de gestão, e.g.: [Sebba Patto 2010] e [Lay et al. 2013].

Percebemos grandes contribuições desses trabalhos. Porém, alguns poderiam ser melhorados em aspectos como a inclusão de comunidades interessadas no processo; outros, em aspectos como o aumento da complexidade do cenário. Nem sempre é conveniente desprezar o conhecimento da população. No seu trabalho, [Lay et al. 2013] mostram que cidadãos comuns podem ter percepções parecidas àquelas de cientistas. Em [Sebba Patto 2010], o autor mostra que ações participativas podem auxiliar gestões pela educação e aumento do conhecimento sobre o domínio, por parte de cidadãos envolvidos.

Neste artigo, apresentamos uma proposta de um agente computacional para simular um ambiente urbano virtual. Este trabalho foi desenvolvido no seio do projeto Assistente Digital para Gestão Pública Participativa (doravante ADGEPA). Um dos principais objetivos deste agente ambiente é suportar simulação dos dados de modo que os dados gerados tenham coerência com o cenário representado.

Primeiramente, apresentamos as motivações, os objetivos e o projeto ADGEPA. Na Seção 3, discutimos simulação multiagentes no projeto, seus objetivos e seu ciclo. Posteriormente, apresentamos a arquitetura preliminar da simulação. Na Seção 4, explicamos nossa proposta do Agente Ambiente - assunto principal deste trabalho. Antes de concluirmos, apresentamos o estado atual do trabalho e o prospecto das atividades.

2. O projeto ADGEPA

2.1. Motivação

Uma das principais motivações do projeto é investigar a possibilidade de auxiliar a administração pública com uso de tecnologias web e participação de cidadãos. Além disso, acreditamos que problemas socioambientais possam ter associações uns com os outros, seja pelo favorecimento ao crescimento ou ao surgimento de problemas. Assim, pretendemos desenvolver um assistente digital para auxiliar em tomadas de decisão em gestão pública. Além de participativo, este processo pode ser mais eficiente que práticas convencionais de gestão pública.

2.2. Objetivos

De maneira mais abstrata, podemos dizer que o objetivo geral do projeto ADGEPA é dar oportunidade aos membros de comunidades para fazer denúncias sobre determinados problemas socioambientais, de forma que contribuam na escolha de aspectos que possam ser resolvidos pelo poder público. O processo de decisão deve ser baseado na perspectiva da população acerca dos problemas socioambientais mais relevantes.

Para isso, pretendemos disponibilizar um aplicativo web responsivo e um módulo para captação de dados ambientais com uso de redes de sensores (o último não será abordado neste artigo). Além de armazenar e compartilhar informações, o ADGEPA propõe que seja feito refinamento dos dados em busca de conhecimento útil. Este processo de busca por informações úteis é conhecido como KDD (*Knowledge Discovery in Database*, que em português significa Descoberta de Conhecimento em Base de Dados) [Cios et al. 2010]. A partir da descoberta de conhecimento em nossos dados, pretendemos dar suporte a decisão a partir de técnicas que considerem diversos aspectos (socioambientais), e.g. análise multicritério.

2.3. Metodologia

Nosso trabalho está dividido em 4 etapas, cada uma descrita em um dos próximos quatro parágrafos. A segunda etapa, enfatizada neste artigo, consiste na simulação da primeira etapa. Na Seção 3, apresentamos detalhes da simulação, da arquitetura e explicamos o relacionamento das etapas 2, 3 e 4 (representados graficamente pela Figura 2).

A princípio, pretendemos fazer coleta de dados acerca de problemas socioambientais em paisagens urbanas por meio de um aplicativo web responsivo e por rede de sensores. Existem 29 indicadores ambientais, 30 sociais e 4 variáveis ambientais coletadas por sensores. Além do aspecto socioambiental, o sistema registra quem faz a denúncia, onde e quando o fato denunciado é percebido.

Haja vista que o tempo para coleta de dados é muito grande, é preciso simular o aplicativo web e seus usuários. Para isso, propomos um modelo de simulação com um agente computacional para representar um ambiente urbano e agentes usuários para representar diferentes usuários da aplicação. Pela simulação, esperamos disponibilizar dados para os agentes mineradores e gestor. É importante destacar que não encontramos nenhum trabalho com dados capazes de representar todos os aspectos socioambientais tratados no nosso trabalho e nem parte deles em quantidade e diversidade suficiente para fazermos nossos experimentos. Detalhes acerca dos agentes e da simulação na Seção 3.

Pelo uso de KDD, buscamos padrões para promover melhor compreensão dos dados envolvidos na pesquisa. Neste projeto, contamos com a colaboração de dois especialistas do domínio que nos ajudam na compreensão da representatividade e na granularidade dos dados; fatores fundamentais em KDD, segundo [Pereira Dos Santos Silva 2004]. Para identificação de padrões, buscamos agrupamentos e associações para descobrir possíveis correlações entre aspectos socioambientais e existência de séries temporais. Assim, usamos a mineração (parte do trabalho de KDD) para ajudar na descoberta de padrões [Cios et al. 2010] e consequentemente, na descoberta de conhecimento.

Para tratar as informações descobertas, modelamos um agente computacional de suporte a decisão. Esse agente representa um gestor artificial que pode ter diferentes perfis (menos ou mais tendenciosos a aspectos sociais ou ambientais). Para sugerir aspectos a serem atacados, o agente elenca os aspectos socioambientais mais associados a outros aspectos. A partir dos mais associados (identificados na fase de mineração), aplicamos métodos matemáticos de análise multicritério para identificar aspectos mais relevantes para a população de uma determinada comunidade. Por razões práticas, não apresentamos detalhes do agente gestor e da sua metodologia de suporte a decisão neste artigo.

3. Simulação multiagentes no ADGEPA

3.1. Justificativa

A princípio, os dados do projeto deveriam ser providos de usuários do aplicativo web e de redes de sensores. Pelo sistema, os usuários podem fazer ou visualizar denúncias de problemas socioambientais em paisagens urbanas com auxílio de mapas. Porém, percebemos que seria necessário muito tempo para coletar dados e obter volume satisfatório para os fins da pesquisa. Além disso, não encontramos e não dispomos de nenhum trabalho ou base de dados capaz de atender a nossa demanda. Para não prejudicar o andamento das atividades de descoberta de conhecimento e de suporte a decisão, decidimos simular

os dados. Para esta etapa, escolhemos sistemas multiagentes (SMA) porque esta técnica facilita a modelagem dos nossos cenários (cidades). Além disso, SMA permite automatizar tarefas humanas [Russell and Norvig 2009]. Dentre as principais características de agentes inteligentes, segundo [de Freitas 2002] e [Wooldridge 2009], destacamos: autonomia (característica de usuários humanos simulados por Agentes Usuários), cooperação (característica desejável entre Agentes Mineradores e Agente Gestor), capacidade de operar em ambientes dinâmicos (embora mapas sejam estáticos, os agentes modelados devem ser capazes de agir em diferentes ambientes), proatividade (características de usuários humanos, simulados pelos Agentes Usuários) e reatividade (característica do ambiente, representado pelo Agente Ambiente).

3.2. Os agentes

No projeto ADGEPA, está previsto a implementação de quatro tipos de agentes computacionais. Dos quatros agentes, apenas dois tipos estão envolvidos na simulação de dados: agentes Usuários e Ambiente.

O Agente Gestor (denominado AG) tem como objetivo dar suporte a escolha de estratégias para mitigar ou eliminar problemas socioambientais. Para isso, o AG percebe saídas geradas por Agentes Mineradores. Em seguida, o AG identifica elementos mais relevantes para atenuar ou eliminar os problemas socioambientais mais denunciados.

Os Agentes Mineradores (denominado AM) têm como objetivo (a) agrupar denúncias de acordo com localidades e momentos e (b) buscar associações entre denúncias e associações entre denúncias e momentos (identificação de séries temporais).

Os Agentes Usuários (denominados AUs) simulam usuários humanos da aplicação e são responsáveis por denúncias, feitas de acordo com as características de cada AU e a percepção do ambiente.

O Agente Ambiente (denominado AA) representa ambientes virtuais para os AUs. O AA informa os AUs acerca do ambiente (uma cidade) sempre que for consultado. Mais detalhes sobre o AA na seção 4. Neste trabalho não apresentamos detalhes dos outros agentes (que ainda estão em fase de modelagem ou implementação).

3.3. A simulação

Para o nosso modelo, algumas características precisam ser definidas em tempo de execução: o intervalo de tempo (expresso em dias) e a quantidade de AUs que populam o cenário. Outras são pré-definidas: a paisagem urbana e elementos do cenário urbano (escolas, hospitais, feiras, teatros, bares, cinemas, locais de trabalho, etc). A percentagem de AUs que frequentam alguns ambientes (e.g. escolas) foi estabelecida de acordo com dados obtidos do Senso IBGE; outras foram estimadas (e.g. feiras, 1 vez por semana). A Figura 1 representa um cenário urbano escolhido para configurar o AA. Neste caso, parte do mapa de Goiânia é definido pelas coordenadas expressas em graus decimais: (-49.272211, -16.693656), (-49.228052, -16.693656), (-49.272211, -16.671375) e (-49.228052, -16.671375). Para cada tipo de elemento (e.g. escolas) é definido um vetor onde há uma coordenada para cada elemento daquele tipo. Após executarmos a simulação, esperamos popular uma base de dados com denúncias sobre problemas socioambientais, momentos que problemas foram percebidos ou denunciados, locais que problemas foram percebidos e responsáveis por denúncias.

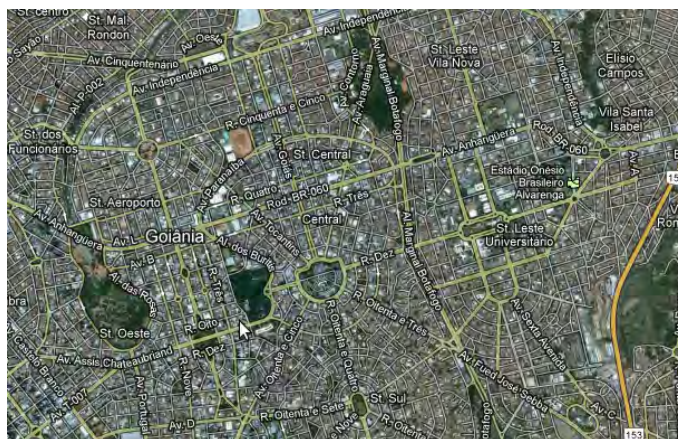


Figure 1. Trecho da cidade de Goiânia

3.4. O ciclo da simulação

A simulação ocorre em ciclos que representam dias. Durante cada ciclo, AUs percorrem rotas (de suas casas ao local de estudo, de trabalho, de entretenimento, etc), percebem o ambiente pela troca de mensagens com o AA e podem fazer marcações (denúncias de problemas socioambientais que ocorrem em determinados locais e momentos). As rotas dos AUs podem variar de acordo com o dia da semana. Por questão de otimização de recursos computacionais, as rotas foram definidas como segmentos de reta entre dois pontos. Embora essa forma de deslocamento não represente fielmente a forma humana convencional de se deslocar, ela não interfere na percepção do ambiente. Os agentes repetem este ciclo (dia) até completar o total de dias da simulação.

Ao término da simulação, os AMs entram em ação. Eles fazem agrupamentos nos dados de acordo com localidades e momentos para buscar associações com intuito de descobrir possíveis correlações entre denúncias e denúncias e momentos.

Por último, o AG faz análise das descobertas feitas pelos AMs e, com base em técnicas matemáticas de análise multicritério, sugere problemas a serem atacados pela administração pública para que se possa resolver ou mitigar a maior parte dos problemas mais relevantes a sociedade.

3.5. A arquitetura preliminar da simulação

Dentre alguns frameworks para construção de sistemas multiagentes podemos citar: Jason [Hübner et al. 2004], Cormas [Le Page et al. 2012] e JADE [Gomes et al. 2003]. Adotamos JADE porque o framework atende satisfatoriamente nossas necessidades, possui bastante documentação e é bem aceito em comunidades científicas sobre simulação multiagentes. Além disso, segue padrões de desenvolvimentos estabelecidos pela FIPA¹.

Na Figura 2, representamos a nossa arquitetura de simulação: instâncias de AUs percebem o ambiente, fazem denúncias e as armazenam na base de dados. Ao concluir o ciclo de simulação, os AMs fazem busca de padrões por agrupamentos e associações. Com base na saída dos AMs, o AG identifica aspectos que podem ser atacados para minimizar ou resolver os problemas socioambientais mais relevantes à população.

¹FIPA - *Foundation for Intelligent Physical Agents*. Disponível em <<http://www.fipa.org/>>

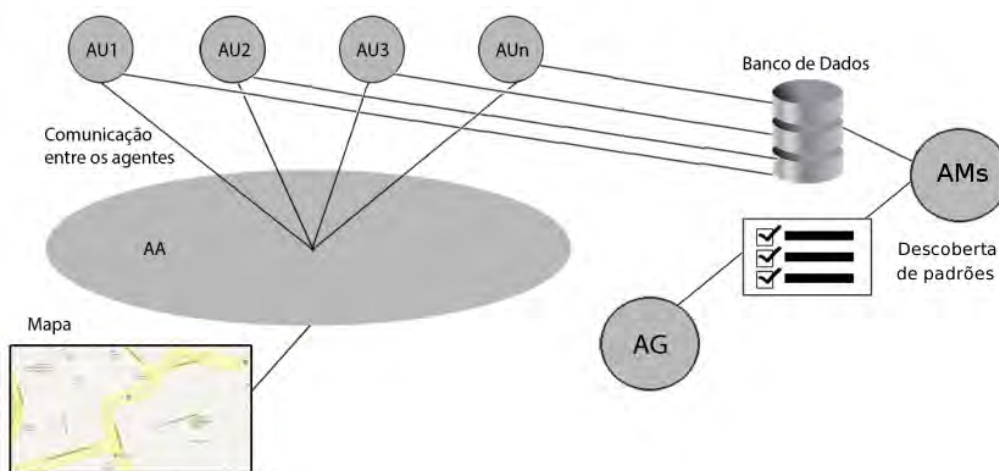


Figure 2. Arquitetura geral da simulação

4. O Agente Ambiente

O AA proposto é reativo, não possui desejos, objetivos e intenções. Sua principal função é representar cenários urbanos de forma que AUs possam compreender os elementos e fazer marcações coerentes com o cenário que eles percorrem.

No início da simulação, o AA é informado sobre a quantidade de dias que a simulação durará e a quantidade de agentes que popularão o cenário. Os elementos do mapa (e.g. feiras e cinemas) têm suas quantidades e coordenadas pré-definidas. Ao serem criados, os AUs conhecem os elementos das suas rotas, mas não sabem onde eles se encontram. No início da simulação, os AUs são informados pelo AA onde se encontram os elementos. Ao se deslocarem, os AUs não sabem o que há abaixo deles, logo precisam perceber o ambiente para que possam fazer denúncias.

Representados em Java, os principais métodos da classe Agente Ambiente são:

- **public void informaElementos()**: informa coordenadas dos elementos da simulação (a princípio, os AUs não sabem onde ficam os elementos de suas rotas).
- **public Coordenada getCoordenada(String)**: retorna a coordenada geográfica do elemento passado como parâmetro. Este método permite que AUs possam alterar dinamicamente suas rotas (e.g. mudar de supermercado).
- **public String oQueE(Coordenada)**: o AA informa qual é o elemento que está em uma determinada coordenada (x, y).

Vale lembrar que em JADE, as mensagens entre agentes não funcionam da mesma forma que em linguagens orientadas a objetos (pela chamada de métodos). Em JADE, cada mensagem tem remetente, lista de destinatários, performativa, conteúdo, linguagem, ontologia, entre outros; detalhes em [Bellifemine et al. 2007].

5. Trabalho atual e prospecto de atividades

Os agentes AUs, AMs e AG se encontram em fase de modelagem. Dos três, o AG se encontra mais adiantado. Já simulamos a aplicação de uma técnica matemática de análise multicritério, os resultados foram bastante satisfatórios. Porém, faremos mais estudos

com outras técnicas antes de publicarmos sobre este agente. O AA está em fase de codificação. Haja vista que os agentes são desenvolvidos por diferentes membros da equipe, será preciso integrá-los futuramente e fazer experimentos para ajustar o modelo.

Ainda é necessário fazer melhorias no AA, haja vista que a API do GoogleMaps (atualmente utilizada) não sabe distinguir todos elementos. Experimentalmente, implementamos métodos para identificar elementos desconhecidos da API, e.g.: bosques e rios. Para isso, marcamos as coordenadas e criamos polígonos para representar bosques e segmentos de retas para representar rios e córregos. Experimentalmente, usamos margem de erro de 5 metros (muito comum em GPS) para decidir se pontos muito próximos a rios e bosques fazem parte ou não deles. Pretendemos tornar o ambiente dinâmico (i.e. modificar características do cenário em tempo de execução, e.g. condições de trânsito).

6. Conclusão

O objetivo do projeto ADGEPA consiste em disponibilizar uma ferramenta web responsiva para registrar denúncias sobre problemas socioambientais em base de dados; a partir de dados, pretendemos descobrir informações; e das informações, indicamos aspectos que são maiores causadores ou facilitadores do surgimento ou aumento de problemas.

Simulação multiagentes foi escolhida para simular usuários do sistema porque experimentos tomam muito tempo para coletar dados em quantidade satisfatória para mineração de dados. Para simular usuários humanos, propomos Agentes Usuários e um Agente Ambiente. Para descoberta de informações, propomos os Agentes Mineradores. Para suporte a decisão, propomos o Agente Gestor. O Agente Ambiente foi apresentado neste trabalho e se encontra em estado de desenvolvimento. Embora o AA não seja um agente cognitivo, ele possui grande importância porque ele permite fazer com que diversos ambientes de simulação possam ser usados pelos AUs.

Destacamos que não pretendemos simular fidedignamente nem usuários humanos, nem comunidades - a complexidade seria muito alta. Temos grandes expectativas acerca das simulações, dos trabalhos de descoberta de conhecimento e de suporte a decisão.

Por último e não menos importante, destacamos que descoberta de conhecimento e suporte a decisão em dados simulados não têm muita valia no mundo real; servem para validar modelos computacionais dos agentes mineradores e gestor. Ao final, com esses dois modelos, pretendemos verificar se é possível dar suporte a uma gestão participativa.

7. Agradecimentos

Agradecemos aos professores Ivanilton José de Oliveira e Patrícia de Araújo Romão e aos alunos Alex Pereira Maranhão, Leonardo Afonso Amorim, Carlos Henrique Barros Costa, James Lima Cipriano Mota por suas contribuições no projeto. Atualmente, o projeto é financiado pela FAPEG.

References

Alves, H. P. d. F. (2006). Vulnerabilidade socioambiental na metrópole paulistana: uma análise sociodemográfica das situações de sobreposição espacial de problemas e riscos sociais e ambientais. *Revista Brasileira de Estudos de População*, 23:43 – 59.

- Barcellos, C., Pustai, A. K., Weber, M. A., and Brito, M. R. V. (2005). Identificação de locais com potencial de transmissão de dengue. *Revista da Sociedade Brasileira de Medicina Tropical*, 38:246–250.
- Bellifemine, F., Caire, G., and Greenwood, D. (2007). *Developing Multi-Agent Systems with JADE*. Wiley, The Atrium, Southern Gate, Chichester, West Sussex PO19 8SQ, England.
- Cios, K. J., Pedrycz, W., Swiniarski, R. W., and Andrzej Kurgan, L. (2010). *Data Mining: A Knowledge Discovery Approach*. Springer, 1st edition.
- de Freitas, F. L. G. (2002). *Sistemas multiagentes cognitivos para recuperação, classificação e extração integradas de informação da web*. Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Engenharia Elétrica.
- Gomes, E. R., Silveira, R. A., and Vicari, R. M. (2003). Utilização de agentes fipa em ambientes para ensino a distância. In *XXIX Conferência Latino Americana de Informática*, La Paz.
- Hübner, J., Bordini, R., and Vieira, R. (2004). Introdução ao desenvolvimento de sistemas multiagentes com jason. In *XII Escola de Informática da SBC*, volume 2, pages 51–89, Guarapuava. UNICENTRO.
- Katukiza, A., Ronteltap, M., Oleja, A., Niwagaba, C., Kansiime, F., and Lens, P. (2010). Selection of sustainable sanitation technologies for urban slums – a case of bwaise iii in kampala, uganda. *Science of The Total Environment*, 409(1):52 – 62.
- Lay, Y.-F. L., Piégay, H., and Rivière-Honegger, A. (2013). Perception of braided river landscapes: Implications for public participation and sustainable management. *Journal of Environmental Management*, 119(0):1 – 12.
- Le Page, C., Becu, N., Bommel, P., and Bousquet, F. (2012). Participatory agent-based simulation for renewable resource management: The role of the cormas simulation platform to nurture a community of practice. *J. Artificial Societies and Social Simulation*, 15(1).
- Malekpour, S., Langeveld, J., Letema, S., Clemens, F., and van Lier, J. B. (2013). Judgment under uncertainty; a probabilistic evaluation framework for decision-making about sanitation systems in low-income countries. *Journal of Environmental Management*, 118(0):106 – 114.
- Pereira Dos Santos Silva, M. (2004). Mineração de dados - conceitos, aplicações e experimentos com weka. *Escola Regional de Informática RJ/ES*, 1:19–21.
- Poch, M., Comas, J., Rodríguez-Roda, I., Sànchez-Marrè, M., and Cortés, U. (2004). Designing and building real environmental decision support systems. *Environmental Modelling and Software*, 19(9):857–873.
- Russell, S. and Norvig, P. (2009). *Artificial Intelligence: A Modern Approach*. Prentice Hall, 3rd edition.
- Sebba Patto, V. (2010). *The use of assistant agents for the analysis and decision making in participatory management*. PhD thesis, UFR d’Informatique, Université Pierre et Marie Curie.
- Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. Wiley, 2nd edition.

FlexTest – Um Framework Flexível para a Automação de Testes

Camila Socolowski^{1, 2}, André Alarcon², André Temple de Antonio²

¹Departamento de Engenharia de Computação e Automação Industrial (DCA)
Universidade Estadual de Campinas (Unicamp) – Campinas, SP – Brasil

²CPqD Telecom & IT Solutions / DSB – Diretoria de Soluções em Billing
Campinas, SP – Brasil.

{camilas, aalarcon, andret}@cpqd.com.br

Abstract. *Regression testing is applied to ensure the quality of software across multiple versions. This paper presents a framework designed to automate tests recorded by Selenium tool and to allow greater flexibility in the maintenance of test cases by test analysts with different skills. Moreover, the framework provides some facilitators resources to be used at the conference of the expected results of the execution of test cases. The feasibility of this framework was validated through its application in automated testing of a real software.*

Resumo. *O teste de regressão é aplicado para garantir a qualidade de software ao longo de várias versões. Este artigo apresenta um framework desenvolvido para automatizar testes gravados por meio da ferramenta Selenium e para permitir uma maior flexibilidade na manutenção dos casos de teste pelos analistas de teste com diferentes perfis. Além disso, o framework disponibiliza alguns recursos facilitadores a serem utilizados na conferência dos resultados esperados da execução de casos de teste. A viabilidade desse framework foi validada por meio de sua aplicação em automação de testes de um software real.*

1. Introdução

Após seu desenvolvimento, um produto de software usado na indústria precisa evoluir devido à inclusão de novas funcionalidades requeridas pelo cliente a medida que seus negócios mudam. Além disso, defeitos são encontrados durante o uso do software em produção, provocando a necessidade de manutenção. Diante dessa evolução, é um desafio manter a qualidade de software após diversas versões [Park et. al. 2008].

O teste de regressão é usado para garantir qualidade de software, após diversas versões de um produto de software terem sido geradas durante seu desenvolvimento e manutenção. No entanto, o teste de regressão é muito caro, porque requer muitas execuções de casos de teste e o número de casos de teste aumenta consideravelmente à medida que o software evolui [Elbaum et al. 2002].

Esse artigo abordará as técnicas de automação relacionadas à tecnologia utilizada no processo de automação. Mas recomenda-se que antes de iniciar qualquer projeto de automação, sejam analisados o custo-benefício da aplicação das técnicas

seleção de testes de regressão (STR) para se obter um resultado mais eficaz na seleção dos casos de teste que serão automatizados.

O processo de automação definido nesse trabalho aponta primeiramente para a necessidade de identificar com critério as funcionalidades que devem ser automatizadas. Após a identificação dessas funcionalidades e a seleção de casos de teste, o próximo passo é identificar quais as tecnologias e ferramentas poderiam ser utilizadas para automatizar os testes selecionados, e dentro desse contexto identificou-se a necessidade da criação de um framework para automatizar os testes.

Portanto, nesse artigo é proposto o framework FlexTest que tem como objetivo possibilitar a execução de testes funcionais baseados em plataforma web. A aplicação do framework é apresentada em um estudo de caso realizado em um sistema considerado de grande porte.

Este artigo está organizado da seguinte forma: na Seção 2 são descritos os conceitos básicos de testes de regressão e técnicas de automação de testes, na Seção 3 é apresentado o processo de automação utilizado neste trabalho, na Seção 4 é apresentado o framework FlexTest para automação de testes de software, na Seção 5 é apresentado um estudo de caso em que o framework é aplicado em um sistema real e finalmente na Seção 6 são apresentados a conclusão e as sugestões de trabalhos futuros.

2. Definições e Conceitos Básicos

Nas subseções a seguir, são apresentados os conceitos básicos relacionados a teste de regressão e técnicas de automação de testes.

2.1. Teste de Regressão

O teste de regressão é usado para apoiar as atividades de teste e garantir a obtenção da qualidade apropriada ao longo de diversas versões de software [Park et al. 2008]. Quando componentes novos ou modificados, inseridos em um software, causam defeitos a outros componentes não modificados, pode-se afirmar que o sistema em teste regrediu [Binder 2000]. Por isso, os testes aplicados ao novo sistema, contendo as alterações, são chamados “testes de regressão”, que visam evitar a regressão do sistema. As principais técnicas para teste de regressão são: seleção de testes de regressão (STR) [Rothermel and Harrold 1996] e priorização de testes de regressão [Elbaum et al. 2002].

2.2. Técnicas de Automação de Testes

As principais técnicas de automação de teste apresentadas na literatura são: *record & playback*, *scripts*, *data-driven* e *keyword-drivers*.

A técnica *record & playback* consiste em utilizar uma ferramenta de automação de testes para gravar as ações executadas pelo usuário, que interage com a interface gráfica do sistema e converte as ações em scripts de teste que podem ser executados quantas vezes forem necessárias. Essa técnica é considerada simples e prática [Fewster 1999] [Fewster 2001] [Hendrickson 1998].

A técnica de *scripts* é considerada como uma extensão da técnica *record & playback*. Através da programação, os scripts de teste gravados são alterados para que desempenhem um comportamento diferente do script original durante sua execução.

Para que esta técnica seja utilizada, é necessário que a ferramenta de gravação de scripts de teste possibilite a edição dos mesmos. Em comparação com a técnica *record & playback*, a técnica de *scripts* apresenta maior taxa de reutilização, maior tempo de vida, melhor manutenção e maior robustez dos scripts de teste [Tervo 2001].

A técnica *Data-Driven* consiste em extrair, dos scripts de teste, os dados de teste, que são específicos por caso de teste, e armazená-los em arquivos separados dos scripts de teste. Os scripts de teste passam a armazenar os procedimentos de teste (lógica de execução) e as ações de teste sobre a aplicação, que normalmente são genéricos para um conjunto de casos de teste [Nagle 2000] [Zambelich 1998].

A técnica *Keyword-Driven* consiste em extrair dos scripts de teste o procedimento de teste que representa a lógica de execução. Os scripts de teste passam a conter apenas as ações específicas de teste sobre a aplicação, as quais são identificadas por palavras-chave. Estas ações de teste são como funções de um programa, podendo inclusive receber parâmetros, que são ativadas pelas palavras-chave a partir da execução de diferentes casos de teste. O procedimento de teste é armazenado em um arquivo separado, na forma de um conjunto ordenado de palavras-chave e respectivos parâmetros [Fewster 1999] [Nagle 2000].

As ferramentas que fornecem suporte aos testes de sistema geralmente integram uma ou mais destas técnicas de automação permitindo sua utilização na execução de vários tipos de testes de sistema, como os testes funcionais, os testes de regressão e os testes de performance e etc.

3. Processo de Automação de Teste

O processo de automação definido e aplicado no CPqD foi dividido em algumas fases como identificação do que deve ser automatizado, definição dos casos testes e resultados esperados e identificação da tecnologia que será utilizada na automação dos casos de teste. O processo de automação foi definido em algumas fases:

3.1. Identificação do escopo da automação

Esta fase define quais são as funcionalidades que devem ser automatizadas. Uma das fases mais importantes no processo de automação é a decisão do que se deve automatizar. Durante a execução dessa fase foram analisados os dados históricos relacionados à abertura de defeitos pelo cliente na homologação e produção do sistema. Alguns passos importantes foram dados nessa etapa:

- Análise, seleção e sumarização dos defeitos de regressão abertos pelos clientes em homologação e/ou em produção por funcionalidade;
- Definição das funcionalidades que serão automatizadas.

3.2. Identificação dos casos de teste

Esta fase define quais os casos de teste devem ser automatizados da funcionalidade que foi selecionada na fase anterior. Nessa fase, foi necessário identificar os casos de teste que poderiam ser do fluxo básico e/ou alternativo, e os pontos de verificação (oráculos)

para cada caso de teste. Um caso de teste pode apresentar mais de um oráculo a ser validado.

3.3. Identificação da tecnologia

Esta fase define a forma como os oráculos definidos na fase anterior serão validados e quais serão as tecnologias que deverão ser utilizadas para essa validação. Nesse passo, foi realizado um estudo de técnicas de automação e ferramentas de automação.

Após análise dos oráculos para cada caso de teste, verificou-se a necessidade de realizar algumas verificações que não eram atendidas pela ferramenta *Selenium*, que utiliza a técnica “*record & playback*”. Com isso, surgiu a necessidade da criação de um framework que permitisse verificações mais complexas, como por exemplo, verificações em banco de dados, que se integrasse com a ferramenta *Selenium*, que fosse mais flexível para inclusão de novos tipos de verificações e que permitisse a manutenção mais fácil e ágil dos casos de testes gravados.

4. Framework FlexTest

4.1. Arquitetura do Framework

O framework FlexTest é apresentado na Figura 1 em um diagrama que exhibe os seus componentes. O núcleo principal do sistema é o Controle de Execução (CE) que é responsável por disparar e controlar todo o processamento de uma planilha de automação de testes. Ao ser iniciado, o CE gera um arquivo de *log* que irá conter as informações dessa execução. O primeiro passo a ser realizado é a identificação dos arquivos do tipo CSV que estão armazenados na pasta de entrada de dados. Para cada arquivo o CE irá criar um log específico e iniciará o *parser* de cada um desses arquivos.

O componente *parser* é capaz de interpretar uma sequência de comandos de um arquivo no formato CSV e preparar os comandos para serem executados pela componente CE. Algumas regras foram criadas para facilitar o preenchimento e entendimento do arquivo como, por exemplo, inclusão de comentários e continuação de um comando na próxima linha, a fim de facilitar a identificação. Ao finalizar o *parser* dos arquivos, os comandos e parâmetros correspondentes de cada arquivo CSV são retornados para o CE.

Um arquivo de configuração irá conter todos os possíveis comandos do framework bem como o respectivo caminho para cada implementação. Caso ocorra qualquer problema no comando identificado pelo *parser*, o CE imediatamente aborta a execução do CSV em processamento. Caso contrário, os passos de validação e execução de cada comando serão disparados. A validação tem como objetivo identificar possíveis problemas de sintaxe nos comandos e parâmetros antes da execução dos comandos no arquivo CSV.

Uma validação não efetiva da sintaxe dos comandos pode permitir que uma eventual falha só apareça durante a execução dos testes automatizados, e isso acarretaria em prejuízo para os testes. Nesse caso é interessante antecipar ao máximo a identificação de problemas nos comandos do arquivo CSV. A execução dos comandos é o último passo a ser gerenciado pelo CE de acordo com a funcionalidade de cada comando.

Uma das características interessantes do framework é a facilidade com que um novo comando pode ser disponibilizado em sua estrutura. Para isso, o novo comando deve ser implementado em uma classe Java, contendo os serviços de validação e execução e incluído no arquivo de configuração. Todas as tarefas de chamadas ao novo comando, *log*, e outras, são atribuídas ao framework e ficam transparentes à implementação dos comandos.

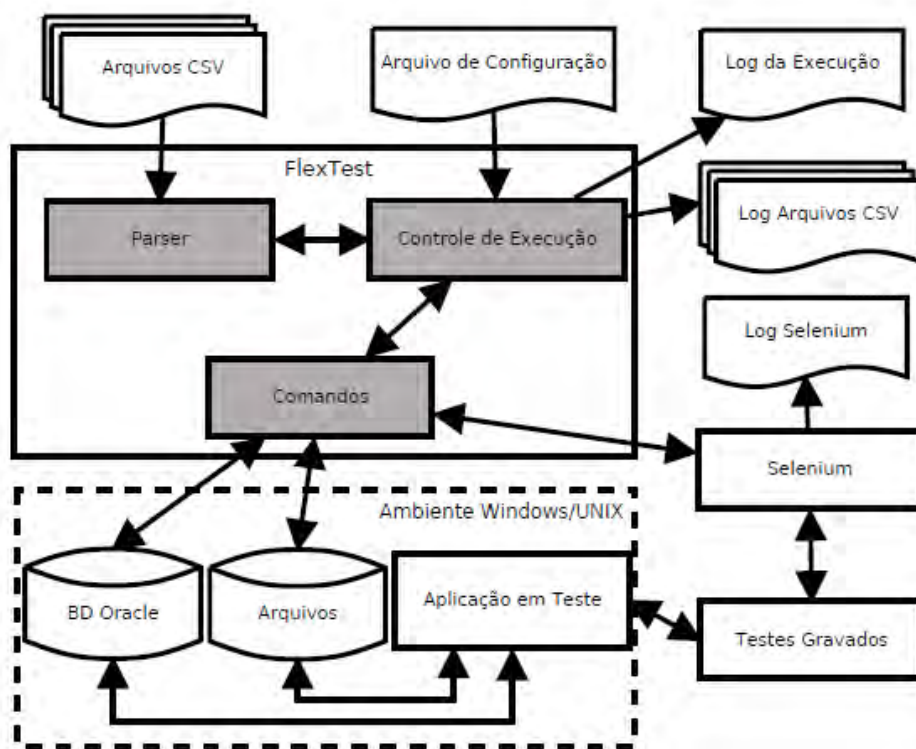


Figura 1 – Arquitetura Completa do Framework FlexTest

Um dos comandos disponibilizados pelo framework executa testes da ferramenta de automação Selenium gerando um *log* específico. Essa ferramenta de automação de teste além de suporte nativo a técnica *'record & playback'*, oferece a possibilidade de exportar os casos de teste gravados para uma linguagem de programação, como por exemplo Java, o que a torna bastante compatível com a arquitetura do framework FlexTest.

O framework possui outros comandos como verificação de SQL, comparação entre arquivos, import e export de banco, etc.

4.2. Funcionamento do Framework

O arquivo CSV que é utilizado como entrada para o framework contém um conjunto de passos (comandos) que serão executados para cada teste. Por exemplo, importar determinado banco de dados, executar um teste específico gravado por meio da ferramenta *Selenium* e executar uma lista de consultas em banco.

Os testes gravados por meio do *Selenium* são portáteis e incrementais. Para adicionar um novo caso de teste ao framework (teste gravado pelo Selenium), o analista de teste deve incluir o código Java gerado pelo Selenium em um arquivo, em um diretório específico que está configurado em sua máquina e clicar no ícone *build*.

O framework disponibiliza também um comando flexível para conferência dos dados no banco. Como todo teste automatizado verifica um resultado esperado, o usuário deverá incluir como parâmetro desse comando as consultas e os resultados esperados para determinado teste. O framework, no momento da execução desse comando, se encarrega de acessar o banco de dados, executar as consultas desejadas e comparar os resultados obtidos com os resultados esperados, incluindo essas informações no *log* de execução do arquivo CSV em teste.

5. Estudo de caso – Utilização do Framework FlexTest na Automação de um Sistema de Gestão de Telecomunicações

5.1. Características do Sistema

A validação do framework foi realizada através da sua aplicação na automação de testes de um sistema de faturamento e de um sistema de gestão de recursos de telecomunicações (GRT) da Fundação CPqD.

O sistema apresenta aproximadamente 300.000 linhas de código e se aplica ao mercado de telecomunicações e sua especificação funcional engloba regras de negócios complexas para empresas operadoras de telecomunicações.

5.2. Fases do Processo de Automação no Sistema de Faturamento e no Sistema GRT

O processo de automação no sistema de faturamento e no GRT passou pelas fases de identificação do escopo, identificação dos casos de teste e identificação das tecnologias descritas na seção 3 desse artigo.

Foram selecionados para iniciar o processo de automação 40 funcionalidades, 250 casos de testes e o framework como ferramenta de automação de testes. As gravações foram realizadas utilizando a ferramenta *Selenium* que possibilitou a exportação do código Java que foi incorporado dentro dos casos de teste do framework.

Após a gravação e a inclusão dos casos de teste no framework, foi criado o arquivo CSV que direcionou a execução dos testes e as conferências que foram realizadas.

5.3. Resultados Obtidos

Nesta seção são apresentados os resultados obtidos na aplicação da automação de teste no sistema de faturamento e no sistema GRT.

Durante todo o processo de automação pode-se observar a facilidade que os analistas de teste com diferentes perfis tiveram para gravar e/ou alterar casos de teste existentes incluindo-os com facilidade na execução do framework. Isso foi possível observar, pois o tempo real gasto nas gravações, alterações de scripts e execuções com sucesso do teste automatizado foi inferior ao tempo estimado para essas mesmas

atividades, em experiências anteriores de automação, quando outras ferramentas eram utilizadas e era necessária a intervenção de um implementador para avançar na automação dos testes.

Foi possível observar a facilidade de verificação de alguns resultados esperados para alguns testes mais complexos de serem validados na automação por meio de comandos SQL que comparam o valor retornado do banco com o valor esperado pelo analista de teste e por fim registram no *log* se o resultado da consulta está correto ou incorreto de acordo com os parâmetros informados.

6. Trabalhos Relacionados

Existem outros trabalhos que abordam ferramentas de suporte à automação de teste de software, porém estas ferramentas apresentam limitações em suas funcionalidades. O framework FlexTest foi criado para atender a uma premissa básica de facilidade, tanto no sentido de ser utilizada por diferentes perfis de analistas de teste, como a de permitir a inclusão de novas funcionalidades com maior agilidade, como também permitir facilidades na verificação de resultados esperados e flexibilidade nas configurações.

Uma grande parte das ferramentas de automação de teste disponíveis são voltadas para a técnica *record & playback*. As ferramentas Functional Tester, WinRunner, QuickTestPro, TestSmith e SilkTest, além da funcionalidade de gravação de scripts de teste, oferecem uma linguagem de programação permitindo que a técnica de programação de scripts seja utilizada.

No trabalho de [Fantinato et. al, 2004] é apresentado um framework que não podia ser utilizado para sistemas baseados na plataforma web, não possibilitava que o processo de automação fosse assumido totalmente pelo analista de teste sem conhecimento em linguagens de programação e solicitava o preenchimento de muitos dados de teste nas planilhas, o que aumentava o tempo de execução dessa atividade.

O framework FlexTest foi criado para fornecer um conjunto maior de suporte a automação de testes. A ferramenta gratuita *Selenium* é um dos componentes deste framework. Um dos objetivos do framework é possibilitar a execução de testes para sistemas baseados em plataforma web.

7. Conclusão e Trabalhos Futuros

Esse trabalho apresenta um framework que foi criado e utilizado no processo de automação do sistema de faturamento da Fundação CPqD para o mercado de telecomunicações. Observou-se que para obter um resultado mais satisfatório na automação dos testes é necessário considerar a tecnologia que será utilizada, mas é necessário também considerar outras questões como a seleção das funcionalidades e dos casos de teste que devem ser automatizados.

Algumas variáveis como o custo alto da automação, manutenção dos casos de teste automatizados, falta de flexibilidade, diferentes perfis de analistas de teste envolvidos no processo de automação, foram consideradas na especificação do framework e na definição das fases do processo de automação.

Como trabalho futuro pretende-se estender o framework FlexTest para realizar comparações entre outros tipos de arquivos, localizar e comparar alguns itens dentro do

arquivo e incluir novos comandos a serem utilizados pelos analistas de teste. Além disso, pretende-se aplicar o processo de automação em outros sistemas e coletar uma quantidade maior de informações relacionadas ao processo de automação.

8. References

- Park H., Ryu, H., and Baik, J.(2008). Historical value-based approach for cost-cognizant test case prioritization to improve the effectiveness of regression testing. *Ssiri*, 0:39-46.
- Elbaum,S., Malishevsky, A.,and Rothermel, G.(2002). Test case prioritization: A family of empirical studies. *IEEE Transactions on Software Engineering*, 28(2):159-182.
- Kim,J.,-M. and Porter, A.(2002). A history-based test prioritization technique for regression testing in resource constrained environments. In *ICSE 2002: Proceedings of the 24th International Conference on Software Engineering*, pages 119-129. New York, NY, USA.ACM.
- Binder, R. V. (2000). *Testing Object-Oriented Systems: Models, Patterns and Tools*. Addison-Wesley Longman.
- Rothermel G. and Harrold, M.J.(1996). Analyzing regression test selection techniques. *IEEE Trans. Softw. Eng.* 22(8): 529-551.
- Fewster, M. & Graham, D., “Software Test Automation”, Addison-Wesley, 1999.
- Fewster, M., “Common Mistakes in Test Automation”, *Proceedings of Fall Test Automation Conference*, 2001.
- Hendrickson, E., “The Differences Between Test Automation Success And Failure”, *Proceedings of STAR West*, 1998.
- Tervo, B., “Standards For Test Automation”, *Proc. of STAR East*, 2001.
- Nagle, C.,“ Test Automation Frameworks”, disponível em <http://members.aol.com/sascanagl/DataDrivenTestAutomationFrameworks.htm>, 2000.
- Zambelich, K., “Totally Data-driven Automated Testing”, disponível em http://www.sqatest.com/w_paper1.html, 1998.
- Fantinato, M., Cunha, A., Dias, S., Mizuno, S., and Cunha, C. (2004). AutoTest–Um Framework Reutilizável para a Automação de Teste Funcional de Software. *Simpósio Brasileiro de Qualidade de Software*.
- CPqD. Disponível em <http://www.cpqd.com.br>.
- Functional Tester. em; <http://www-01.ibm.com/software/awdtools/tester/functional>.

Convolução como Caso de Estudo para Formalização em PVS da Correção Funcional de Implementações em FPGA's

Ariane Alves Almeida¹, Jones Yudi Mori², Mauricio Ayala-Rincón^{1,3}

¹Departamento de Ciência da Computação – Universidade de Brasília (UnB)

²Departamento de Engenharia Mecânica – Universidade de Brasília (UnB)

³Departamento de Matemática – Universidade de Brasília (UnB)

arianealves@aluno.unb.br, {ayala, jonesyudi}@unb.br

Resumo. *É apresentado um caso de estudo preliminar para verificação da correção funcional de operações algébricas implementadas em FPGA's. Tal caso é acerca de filtros de imagens baseados em convolução para tratamento de imagens digitais. A metodologia utilizada é baseada na equivalência funcional entre a definição matemática das operações e seu respectivo operador implementado em hardware. Foi utilizado o PVS como assistente de prova para facilitar a verificação proposta.*

1. Introdução

Com o objetivo de fornecer garantia de funcionamento adequado de um sistema, pode-se utilizar técnicas que o validem de modo sistemático, seja ele implementado em *software* ou *hardware*. Uma das formas de se obter tal garantia é por meio da verificação formal, que será aqui discutida.

Assim, o intuito deste trabalho é apresentar um caso de estudo que ilustra a Verificação por Equivalência Funcional como forma de verificar formalmente as propriedades funcionais de sistemas implementados em *hardware* com dispositivos reconfiguráveis. Para tal, utilizar-se-á o assistente de demonstração *Prototype Verification System* (PVS), a convolução para filtros de imagens e sua implementação em *Field Programmable Gate Arrays* (FPGA's).

A Seção 2 apresenta os conceitos referentes às arquiteturas reconfiguráveis e as arquiteturas FPGA; a Seção 3 as noções de formalização de sistemas e o assistente PVS; a Seção 4 explica como aplicar técnicas de verificação formal para comprovar matematicamente a correção funcional de sistemas implementados em FPGA's; a Seção 5 introduz um caso de estudo simples de verificação de um sistema implementado em FPGA; finalmente, a Seção 6 apresenta conclusões e trabalho futuro.

2. Sistemas Embarcados

Atualmente sistemas computacionais são concebidos através das mais diversas abordagens de *software*, *hardware* e *software-hardware*. Podendo estas duas últimas serem feitas através de dispositivos eletrônicos de propósitos específicos e arquitetura fixa, os *Application-Specific Integrated Circuits* (ASIC's), e também por abordagens híbridas que apresentam tanto características de *software* quanto de *hardware*, que é o caso dos dispositivos programáveis de propósito geral de arquitetura reconfigurável, como FPGAs ou *Complex Programmable Logic Device* (CPLD's) [Gaj and Chodowiec 2009].

2.1. FPGAs

Constituídos basicamente de blocos lógicos, registradores, blocos de entrada/saída e de tabelas de verdade, os FPGA's possibilitam sua programação através de linguagens de descrição de *hardware*, (similares às linguagens de programação de *software*). Isso fornece um nível de abstração muito maior ao projetista do que manipular diretamente elementos físicos para criar seu sistema e facilita o processo de desenho do circuito.

Um projeto em FPGA fornece uma velocidade muito maior quando comparado ao mesmo projeto concebido em *software* e executado em um processador de propósito geral, pois permite uma execução não sequencial e pode ser configurado para conter apenas as operações necessárias à execução do procedimento nele implantado, enquanto um circuito de propósito geral com instruções fixas, mesmo tendo um custo de produção mais baixo, necessita de todas as possíveis operações para qualquer estrutura de dados em seu projeto [Huffmire et al. 2010]. Outras vantagens observadas são o tempo e o custo de produção, que são bem inferiores aos da criação de um ASIC, e a possibilidade de reuso do circuito.

Ao finalizar um projeto em FPGA, é necessário verificar se suas funções estão realmente sendo realizadas e se estão executando corretamente, ou seja, se o resultado obtido é o esperado. Então, após a criação do sistema, é necessária sua validação.

Tal validação é obtida de diversas maneiras, sendo mais comumente utilizados os **testes funcionais**, que são simulações com diversos dados de entrada e comparação com os dados de saída esperados caso as funções tenham sido implementadas corretamente, e os **testes estruturais** ou **análise de timing**, que levam em conta a estrutura interna do sistema criado e o *hardware* em que será inserido, como as conexões criadas e os possíveis atrasos dos sinais em determinada placa de *hardware*. Outra técnica menos difundida, mas que vem ganhando cada vez mais espaço e será melhor discutida na próxima Seção, é a **verificação formal** dos circuitos criados. Esse método pode ser aplicado em diversos estágios do desenvolvimento tanto de *software* quanto de *hardware* e sua aplicabilidade em sistemas FPGA será explicada na Seção 5.

3. Formalização de Sistemas

A verificação formal visa, por meio da aplicação de técnicas baseadas em formalismos matemáticos e lógicos, certificar que determinadas funcionalidades e propriedades dos sistemas e dados processados por estes sejam atendidas em qualquer situação. Além de ser uma abordagem matemática que visa especificar um sistema de maneira simples, completa e precisa para fornecer uma comunicação sem margens para ambiguidade aos projetistas, é uma tarefa muito complexa, bastante minuciosa e exaustiva, o que faz com que sua realização demande muito tempo.

A formalização se dá por meio da especificação do sistema e de provas de correção da sua funcionalidade. A fim de facilitar tal tarefa e evitar que sejam introduzidos erros oriundos de falha humana, foram criados os provadores de teorema, ou assistentes de prova, dos quais se destacam ACL2, Coq, HOL e PVS [Bernardo and Cimatti 2006].

3.1. PVS

Dentre diversos ambientes de especificação e formalização, foi escolhido o PVS, que inclui uma linguagem de especificação funcional com um sistema de tipos elaborado que

permite polimorfismo, subtipagem e tipos dependentes, e que tem uma linguagem de prova *à la Gentzen* com um sistema dedutivo de ordem superior, uma linguagem de especificação, um analisador léxico, um chegador de tipos, bibliotecas de especificação e várias ferramentas de navegação, fornecendo um ambiente integrado para especificação e desenvolvimento de provas formais [Owre et al. 2001b].

A linguagem de especificação do sistema, mesmo não possuindo diversos recursos das linguagens de programação mais comuns, como laços de repetição e estruturas de dados sofisticadas, possibilita expressar em forma de conjecturas quais propriedades se espera que a especificação cumpra. Para isso, é necessário que quem está especificando um sistema tenha pleno conhecimento deste para conseguir representá-lo totalmente apenas com os recursos do PVS, que são seus tipos primitivos, a partir dos quais é possível criar outros, suas operações, estruturas condicionais e chamadas recursivas.

4. Formalização de Implementações em FPGA

A complexidade dos circuitos criados, sejam ASIC's ou FPGAs, tem aumentado a cada dia, ampliando a importância da verificação de seu correto funcionamento, o que chega a representar 80% do custo do projeto de um sistema [Drechsler 2004]. Assim sendo, vem aumentando o interesse nas técnicas de verificação formal, já que podem garantir correteza funcional de sistemas, o que não é possível com a tradicional análise dos resultados via testes e simulações [Perry and Foster 2005]. Por isso, é uma tarefa justificável e viável apesar do tempo e esforço demandados, especialmente para criação de sistemas críticos [Bernardo and Cimatti 2006].

4.1. Verificação de Equivalência Funcional

Existem basicamente duas maneiras de verificar formalmente um sistema, por checagem de equivalência (*Equivalence Checking* - EC) e por checagem de propriedades (*Property Checking* - PC). Com o método PC são identificadas e especificadas propriedades referentes ao circuito pretendido que são diretamente demonstradas ou testadas via técnicas como *Model Checking*. Tais propriedades então devem ser provadas satisfeitas em todas as possíveis circunstâncias do sistema [Kropf 1997], [Li and Thornton 2010].

Já a verificação por EC é uma técnica mais simples, que objetiva garantir a equivalência entre dois circuitos, que podem estar descritos de maneiras e níveis diferentes de abstração. Para isso, um dos circuitos a ser comparado deve ter sido previamente provado correto (por PC, por exemplo). Neste caso, deve-se estabelecer uma correspondência entre os dois e verificar então sua equivalência [Drechsler 2004], [Li and Thornton 2010].

Para este trabalho escolheu-se trabalhar com EC, supondo como circuito correto a especificação da definição matemática da função implementada sobre os FPGA's. Dessa maneira, assumindo a correção da especificação da definição matemática, pode-se analisar a arquitetura implementada comparando funcionalmente sua especificação com a da definição matemática. A técnica pode ser resumida como a seguir:

- Ambas as arquiteturas a serem comparadas devem ser especificadas de modo a deixar explícito e de maneira conservativa o que é funcionalmente e estruturalmente realizado sobre o FPGA em cada caso.
- Formaliza-se então um teorema de que as especificações são funcionalmente equivalentes, isto é, as mesmas entradas possuem saídas equivalentes. Sendo necessária para tal uma correspondência entre as entradas e objetos por elas manipulados.

5. Caso de Estudo Preliminar

Para ilustrar a aplicação da metodologia EC em PVS com implementação de FPGA, serão apresentadas partes de uma especificação criada para uma arquitetura de filtros de imagens baseados em convolução. Para tanto, serão apresentados alguns conceitos básicos dessa aplicação, bem como uma ilustração da arquitetura básica implementada em FPGA.

5.1. Filtros de Imagem Baseados em Convolução

A grande maioria das aplicações em FPGA's utilizam diversos operadores aritméticos para realizar suas funcionalidades. Dentre eles, no processamento de imagens digitais são utilizados filtros de imagem baseados em convolução, que ao serem desenvolvidos em FPGA's exploram o paralelismo a fim de melhorar o desempenho dessas aplicações.

A operação de convolução envolve uma imagem e os coeficientes do filtro a ser utilizado, assim sendo, dada uma imagem I com M linhas e N colunas, e um filtro (*kernel*) K , com $m \leq M$ linhas e $n \leq N$ colunas, a convolução entre I e K é definida pela Equação 1 [Pedrini and Schwartz 2007], onde $(K \otimes I)(x, y)$ representa o (x, y) -pixel da saída da aplicação do filtro, para $\frac{m}{2} \leq x \leq (M-1) - \frac{m}{2}$ e $\frac{n}{2} \leq y \leq (N-1) - \frac{n}{2}$. O filtro deve ter um tamanho tal que m e n sejam ímpares, de forma que o pixel resultante da aplicação do filtro seja aquele situado ao meio da vizinhança do filtro [Mori et al. 2011]. Logo, ao aplicar em toda a imagem esta equação, teremos como saída a imagem $(K \otimes I)$, que terá tamanho $M - m + 1 \times N - n + 1$.

$$(K \otimes I)(x, y) = \sum_{l=-\frac{m}{2}}^{\frac{m}{2}} \sum_{j=-\frac{n}{2}}^{\frac{n}{2}} K(l, j) * I(x + l, y + j) \quad (1)$$

As máscaras usadas como base para este trabalho são de tamanho fixo 3×3 (Equação 2), o que permite simplificar Equação 1, obtendo a Equação 3.

$$K = \begin{bmatrix} k1 & k2 & k3 \\ k4 & k5 & k6 \\ k7 & k8 & k9 \end{bmatrix} \quad (2)$$

$$(K \otimes I)(x, y) = \sum_{l=-1}^1 \sum_{j=-1}^1 K(l, j) * I(x + l, y + j) \quad (3)$$

A arquitetura básica da convolução aplicando um filtro de tamanho 3×3 , mostrada na Figura 1, pode ser encontrada facilmente na literatura, como por exemplo em [Mori et al. 2011]. Os pixels representados em vermelho ilustram a vizinhança 3×3 na qual será aplicada a máscara, os amarelos ilustram um buffer para armazenamento temporário dos pixels a frente nas primeira e segunda linhas da imagem de entrada e os azuis representam o restante da imagem. O filtro é representado pelos pixels coloridos com os índices ki e o resultado pela saída do registrador verde com o símbolo de somatório.

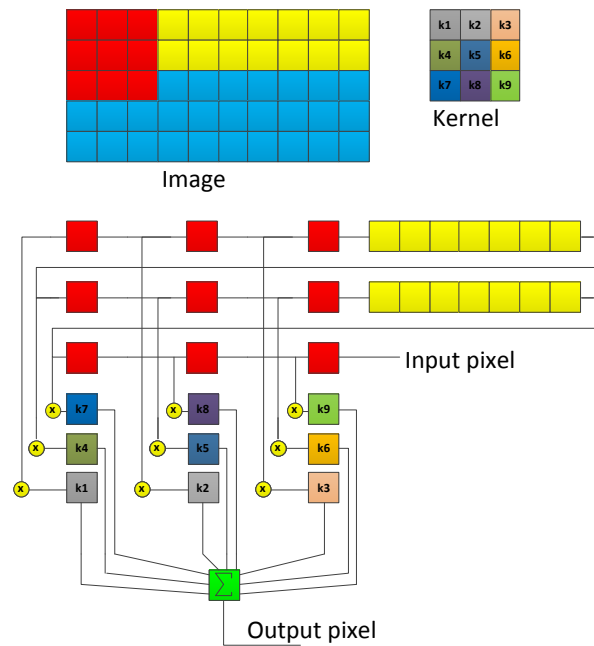


Figura 1. Arquitetura básica para uma convolução 3×3 por [Mori et al. 2011].

Vale ressaltar que a aplicação da definição matemática da convolução tem como resultado uma imagem $K \otimes I$ com $M - 2$ linhas e $N - 2$ colunas, ou seja, a imagem convoluída fica com 1 pixel a menos de borda que a imagem original. Por outro lado, a aplicação da convolução pela arquitetura apresentada gera um stream que representa uma imagem $K \otimes I$ com tamanho $M * N - 2 * N - 2$, ou seja, resulta em alguns pixels distorcidos a mais nas bordas laterais da imagem convoluída que devem ser desprezados.

5.2. Formalização da Arquitetura da Convolução

A fase de prova via EC entre a arquitetura da convolução para filtros de imagens em FPGA e sua equação matemática apresentadas na Seção 5.1 ainda está em andamento, logo, será apresentada apenas a parte de sua especificação em PVS segundo a metodologia da Seção 4 e os artifícios usados para tal.

A função especificada no Algoritmo 1 é relativa à aplicação da convolução em um pixel da imagem com o filtro em uma imagem de tamanho maior ou igual ao deste, e como estão sendo utilizados apenas filtros de tamanho 3×3 , cada índice da matriz K é representado por uma constante ($k1$ a $k9$). Cada imagem `image` é representada por uma sequência finita de linhas de tamanho ≥ 3 (`rows`), e cada linha por sua vez é representada por uma sequência finita de inteiros de tamanho também ≥ 3 (`columns`).

Algoritmo 1 Função da Convolução Matemática em Um Pixel da Imagem.

```
loc_conv(im : image, i: posnat | i < rows - 1,
j : posnat | j < columns - 1) : int =
  im(i-1)(j-1)*k1 + im(i-1)(j)*k2 + im(i-1)(j+1)*k3 +
  im(i)(j-1)*k4 + im(i)(j)*k5 + im(i)(j+1)*k6 +
  im(i+1)(j-1)*k7 + im(i+1)(j)*k8 + im(i+1)(j+1)*k9
```

Esta função é então chamada pela função apresentada no Algoritmo 2, referente à aplicação da convolução, dada na Equação (3), a todos os pixels de entrada da imagem I , que retorna um elemento do tipo `conv`, que é uma sequência finita de linhas de tamanho $rows - 2$, e cada linha é uma sequência finita de inteiros de tamanho $columns - 2$, ou seja, o resultado final de $(K \otimes I)$.

Algoritmo 2 Função da Convolução Matemática em Toda uma Imagem.

```
convolution(im) : conv =
  (# length := rows - 2,
    seq := (LAMBDA (i: below[rows - 2]):
      (# length := columns - 2,
        seq := (LAMBDA (j : below[columns - 2]):
          loc_conv(im, i + 1, j + 1) ) #) ) #)
```

Alguns artifícios foram necessários para poder especificar a arquitetura da convolução. Neste caso, foi utilizada como entrada uma única sequência finita de inteiros `stream` de tamanho ≥ 9 para representar toda a imagem, já que é assim que esta é recebida pela arquitetura da Figura 1. Então, a manipulação de índices foi feita para simular uma matriz de pixels (imagem) e a passagem de todos seus pixels pelo buffer e sua posterior alocação na área vermelha, que é efetivamente a área de aplicação do filtro.

Respectivamente, nos Algoritmos 3 e 4, são especificadas a aplicação local da convolução em apenas uma vizinhança 3×3 descrita pela arquitetura e a aplicação da convolução em toda a imagem, utilizando-se da concatenação de sua chamada recursiva com a convolução local para chegar ao resultado final de $(K \otimes I)$.

Algoritmo 3 Função da Convolução Local em Um Pixel da Imagem pela Arquitetura

```
local_conv_arch(s: stream, (n: numcolumns | n <= s'length),
  (i: nat | i <= s'length - (n * 2) - 3)) : int =
  s(i)*k1      + s(i+1)*k2      + s(i+2)*k3 +
  s(i+n)*k4    + s(i+n+1)*k5    + s(i+n+2)*k6 +
  s(i+(n*2))*k7 + s(i+(n*2)+1)*k8 + s(i+(n*2)+2)*k9
```

Algoritmo 4 Função de Aplicação da Convolução a Todos os Pixels de uma Imagem

```
convolution_arc(s: stream, n: numcolumns,
  (i: nat | i <= s'length - (n * 2) - 3)):
RECURSIVE finite_sequence[int] =
  IF i = s'length - ((n * 2) + 3) THEN
    add_first(local_conv_arch(s, n, i), empty_seq)
  ELSE
    add_first(local_conv_arch(s, n, i), convolution_arc(s, n, i+1))
  ENDIF
MEASURE (s'length - i)
```

Para formalizar a equivalência funcional entre a arquitetura e a equação, é necessário que haja uma correspondência entre suas entradas, como visto na Seção 4.1. Como os elementos de entrada da definição matemática (sequência de sequências) e da arquitetura (sequência de inteiros) são diferentes, os elementos de entrada da primeira são transformados naqueles aceitos como entrada pela segunda através da função especificada no Algoritmo 5 com uso de chamadas recursivas.

Algoritmo 5 Função que Transforma uma Imagem (Sequência de Sequências) em uma Sequência Única.

```
image_to_seq (im: image, i: below[im'length - 2]): RECURSIVE stream =
  IF i = im'length - 3 THEN
    im(im'length - 3) o im(im'length - 2) o im(im'length - 1)
  ELSE
    im(i) o image_to_seq(im, i + 1)
  ENDIF
MEASURE (im'length - i)
```

A partir disso pode-se especificar um lema que expresse a equivalência funcional entre a equação da convolução e a arquitetura ilustrada na Figura 1, através da correspondência entre os elementos do resultado de quando aplicadas a equação e a implementação da arquitetura, sendo necessário desprezar os elementos da borda da imagem gerada pela arquitetura, pois seus resultados apresentam uma borda extra distorcida em relação ao resultado gerado pela equação matemática, como mencionado anteriormente.

O lema do Algoritmo 6 expressa essa correspondência entre os elementos resultantes das duas especificações. Para demonstrá-lo é necessária uma prova por indução aninhada sobre os índices i e j , pois ao percorrer a imagem im , tanto para aplicação das funções quanto para relacionar os elementos dos resultados, estes índices se incrementam, decrementando o tamanho da imagem restante a ser analisada.

Algoritmo 6 Lemma da Equivalência Funcional da Convolução Matemática com sua Arquitetura em FPGA

```
conv_arc_is_same_math : LEMMA
  FORALL (im: image, i: below[rows - 2], j: below[columns - 2]) :
    convolution(im)(i)(j) =
      convolution_arc(image_to_seq(im, 0), columns, 0)(i * columns + j)
```

Além disso, é necessário o uso de lemas auxiliares que relacionem previamente os elementos de uma imagem qualquer com seu stream resultante da aplicação da função `image_to_seq` especificada acima e que determinem exatamente o tamanho do resultado da arquitetura da convolução aplicada à um stream, já que isso não é explícito na definição da função e é de extrema importância durante a prova.

Quando concluída a prova do lema acima, dado que a equação matemática está correta por hipótese, ficará provado que esta arquitetura também é funcionalmente correta.

6. Conclusão

Foi apresentada a importância da verificação formal para garantir a corretude em sistemas implantados em FPGA, que podem ser utilizados como protótipo para a criação de ASIC's corretos, já que estes últimos são mais caros para se produzir e qualquer erro de projeto pode aumentar ainda mais este custo.

O caso de estudo aqui apresentado, além de ilustrar o uso de EC para formalização de sistemas com operadores matemáticos implementados em FPGA's, explana a diferença entre as especificações de uma definição matemática e de sua implementação em um FPGA, mostrando a necessidade do uso de artifícios e manipulações para relacionar os resultados obtidos pelos dois e provar então a corretude funcional dessa arquitetura.

Assim sendo, pode-se utilizar a prova de corretude por equivalência funcional do sistema aqui apresentado, bem como de outros operadores aritméticos para criação e verificação de outros sistemas mais complexos que tenham tais funcionalidades definidas como parte do circuito, a fim de garantir sua corretude.

Referências

- Bernardo, M. and Cimatti, A. (2006). *Formal Methods for Hardware Verification: 6th International School on Formal Methods for the Design of Computer, Communication, and Software Systems, SFM 2006, Bertinoro, Italy, 2006, Advances Lectures*. Lecture Notes in Computer Science / Programming and Software Engineering. Springer.
- Drechsler, R. (2004). *Advanced Formal Verification*. Falk Symposium Series. Springer.
- Gaj, K. and Chodowicz, P. (2009). FPGA and ASIC Implementations of AES. In Çetin Kaya Koç, editor, *Cryptographic Engineering*. Springer, New York, NY - Estados Unidos.
- Huffmire, T., Nguyen, C. I. T. D., Levin, T., Kastner, R., and Sherwood, T. (2010). *Handbook of FPGA Design Security*. Springer.
- Kropf, T. (1997). *Formal Hardware Verification: Methods and Systems in Comparison*. Lecture notes in computer science. Springer.
- Li, L. and Thornton, M. (2010). *Digital System Verification: A Combined Formal Methods and Simulation Framework*. Synthesis Lectures on Digital Circuits and Systems Series. Morgan & Claypool.
- Mori, J., Sanchez-Ferreira, C., Muñoz, D., Llanos, C., and Berger, P. (2011). An unified approach for convolution-based image filtering on reconfigurable systems. In *Programmable Logic (SPL), 2011 VII Southern Conference on*, pages 63 –68.
- Owre, S., Shankar, N., Rushby, J. M., and Stringer-Calvert, D. W. J. (2001a). *PVS Language Reference*. NASA.
- Owre, S., Shankar, N., Rushby, J. M., and Stringer-Calvert, D. W. J. (2001b). *PVS System Guide*. NASA.
- Pedrini, H. and Schwartz, W. R. (2007). *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*. Editora Thomson Learning.
- Perry, D. and Foster, H. (2005). *Applied Formal Verification: For Digital Circuit Design*. McGraw-Hill Electronic engineering. McGraw-Hill Companies, Incorporated.

Mecanismo de Controle de Fluxo Aplicado na Transmissão de Vídeo no Sistema LTE

Igor M. B. Dias¹, Marcelo Palheta¹, Ananias P. Neto¹

¹Centro de Ciências Exatas e Tecnologia (CCET) – Curso de Bacharelado em Ciência da Computação – Universidade da Amazônia (UNAMA)
CEP 66060-902 – Belém – PA – Brasil

dias.bcc@gmail.com, marcellopalheta@gmail.com, apneto@unama.br

Abstract. *This work reports a comparative study of metrics for Quality of Experience (QoE) in LTE system (Long Term Evolution) for different values of GoP (Group of Pictures), using flow control and QoS (Quality of Service) mapping between the eNB (E-UTRAN Node B) and the UE (User Equipment), based on the search of the optimal solution mechanism QFC-LTE (QoE Flow Control LTE). The comparisons made in this paper, discuss the metrics: PSNR (Peak Signal to Noise Ratio), SSIM (Structural Similarity Index) and VQM (Video Quality Metric). The results obtained from the simulations show the impact and the benefits of proposal compared to conventional methods.*

Resumo. *Este trabalho relata um estudo comparativo de métricas de Qualidade de Experiência (QoE) em sistema LTE (Long Term Evolution) para diferentes valores de GoP (Group of Pictures), utilizando o controle de fluxo e mapeamento de QoS (Quality of Service) entre a eNB (E-UTRAN Node B) e o UE (User Equipment), baseado na busca de solução ótima do mecanismo QFC-LTE (QoE Flow Control LTE). As comparações realizadas neste trabalho abordam as métricas: PSNR (Peak Signal to Noise Ratio), VQM (Video Quality Metric) e SSIM (Structural Similarity Index). Os resultados obtidos a partir das simulações mostram o impacto e os benefícios da proposta em comparação aos métodos convencionais.*

1. Introdução

Atualmente, há uma gama de serviços de telecomunicações que transmitem voz, vídeo e dados através de complexas redes de transmissão, sendo que em alguns casos, o serviço não tem uma qualidade aceitável pelo usuário final [Begazo *et al* 2011].

Nas últimas décadas, as redes móveis têm permitido grandes avanços e mudanças nos sistemas de comunicação, garantindo maior mobilidade e disseminação de conteúdo. A Tecnologia LTE (*Long Term Evolution*) é a evolução da tecnologia de rede móvel padronizada pelo 3GPP (*3rd Generation Partnership Project*), que utiliza novos sistemas de múltiplo acesso a interface aérea, para prover novos serviços e aplicações de conteúdo, tais como, MMOG (*Multimedia Online Gaming*), TV Móvel, Web 2.0, *streaming* de vídeo, download de músicas e muitos outros [McQueen 2009].

As redes de celular de próxima geração foram criadas para facilitar o tráfego de comunicações de dados em alta velocidade, além das chamadas de voz. Novos padrões e tecnologias estão sendo implementadas para permitir que redes sem fio substituam linhas de fibra óptica ou cobre entre pontos fixos afastados por vários quilômetros (acesso sem fio fixo) [Rappaport 2009].

Um dos desafios na concepção das redes de próxima geração consiste no desenvolvimento de arquiteturas que viabilizem tanto a continuidade do serviço quanto a visão centrada no usuário, por meio do suporte adequado na QoE para aplicações multimídia, e que proporcionem sempre a melhor conectividade aos usuários móveis [Silva *et al* 2010].

A proposta deste trabalho é avaliar o desempenho da tecnologia de rede móvel LTE, baseada em métricas de QoE na disseminação de conteúdo multimídia. Avaliar a transferência de vídeos para diferentes valores de congestionamento e GoP, utilizando o controle de fluxo e mapeamento de QoS na rede de acesso LTE, compreendido entre a eNB (*E-UTRAN Node B*) e o UE (*User Equipment*). O método utiliza, dentre outros critérios, parâmetros da aplicação de Vídeo, fluxo de dados FTP e VoIP, além da relação do número de UEs e as condições de congestionamento do sistema LTE.

O artigo está organizado da seguinte forma: Na seção 2 são discutidos os trabalhos relacionados. A seção 3 apresenta uma abordagem sobre a rede LTE. As métricas de Qualidade de Experiência são mostradas na seção 4. A seção 5 apresenta o mecanismo de controle de fluxo baseado na formulação matemática do QFC-LTE. As simulações e resultados finais são descritos na seção 6, enquanto a seção 7 apresenta as conclusões e trabalhos futuros.

2. Trabalhos Relacionados

Nesta seção são apresentadas algumas propostas de mecanismos de controle de fluxo, a fim de promover uma melhor qualidade do tráfego de vídeo, levando em consideração as características deste tipo de fluxo.

Em [Begazo *et al* 2011] são realizados testes com um cenário de simulação de rede, onde são parametrizadas diferentes condições de rede, considerando um serviço de vídeo *streaming*, para o qual são consideradas diferentes qualidades de vídeo, utilizando controle de fluxo convencional. Os resultados mostram que a qualidade de vídeo é afetada diretamente pelos fatores como, o atraso e a perda de pacotes.

A proposta apresentada em [Hong *et al* 2010], desenvolveu um método que realiza a transmissão do fluxo de vídeo, denominado SAPS (*Significance Aware Packet Scheduling*), que ajusta os intervalos de tempo entre os pacotes baseado na significância do conteúdo. O trabalho compara os resultados com algoritmos utilizados na literatura, ambos com controle de fluxo *Drop Tail*. O SAPS apresentou uma melhor significância na QoE, quando comparado aos métodos tradicionais.

No trabalho desenvolvido em [Qiu *et al* 2010] é apresentado um controle de fluxo baseado na HVQ (*Hierarchy Virtual Queue*), o mecanismo propõe uma hierarquia de fila virtual evitando a perda de pacotes no sistema LTE. Além disso, o artigo investiga o mecanismo HVQ para diferentes classes de tráfego como, VoIP, Vídeo, *Best Effort*. Os resultados obtidos mostram que o mecanismo HVQ realiza o controle de fluxo com uma quantidade baixa de recursos, no entanto, produz melhor desempenho para o sistema LTE.

3. Rede LTE

A rede LTE é um padrão de tecnologia de banda larga móvel padronizada pela 3GPP, e está sendo desenvolvida para ser utilizada pelas operadoras de redes móveis como evolução das redes 3G (Terceira Geração), resultando, assim, nas Redes 4G (Quarta Geração) e foi projetada para suportar de forma eficiente vários tipos de serviços, em especial orientados a pacotes como, por exemplo, o VoIP. A arquitetura LTE está dividida, segundo [Holma e Toskala 2009], em quatro principais domínios de alto nível:

UE (*User Equipment*), E-UTRAN (*Evolved UTRAN*), EPC (*Evolved Packet Core*) e *Services*.

4. Métricas de QoE

O PSNR é uma métrica objetiva que compara a qualidade do vídeo recebido pelo usuário em relação ao vídeo original. O PSNR, desta forma, determina o valor máximo de luminosidade para cada pixel, comparando a taxa de erro do vídeo recebido. É formulado matematicamente a partir da equação (1), para um quadro original codificado a 8 bpp (256 níveis de cinza) e expresso em dB [Rodríguez e Bressan 2012]:

$$PSNR = 10 \log_{10} \left[\frac{255^2}{MSE} \right] \quad (1)$$

Conforme cita [Rodríguez e Bressan 2012], a métrica SSIM, diferentemente do PSNR, que somente compara a taxa de erro do vídeo recebido em relação ao vídeo original, avalia o vídeo recebido considerando outros fatores, como o HVS (*Human Visual System*). Surgindo em função de o HVS ser altamente eficiente em extrair informações visuais das imagens/vídeos e não das taxas de erro, o SSIM analisa a similaridade, luminosidade, contraste e estrutura. Os valores extraídos do *frame* reconstituído pelo usuário e do *frame* original são armazenados em vetores, separadamente, sendo um vetor para cada cor. Posteriormente, obtém-se a média de cada vetor, e a combinação dessas três médias gera o valor do SSIM, indicando a qualidade do vídeo [Xinbo *et al* 2009]. O valor do SSIM é obtido pela equação (2):

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (2)$$

onde apresenta a função de comparação de brilho na primeira parcela da equação, μ representa a intensidade de brilho médio das imagens x e y , a função comparação contraste é representado pela segunda parcela da equação, onde σ representa o desvio padrão da amostra do brilho das duas imagens x e y , além da função de comparação de estrutura que relaciona a covariância de amostras de luminosidade dos quadros.

A métrica VQM, assim como as demais, baseia-se na comparação do vídeo recebido pelo usuário em relação ao vídeo original, comparando o brilho e o contraste. Conforme afirma [Gelonch *et al* 2010].

5. Método de Controle de Fluxo QFC-LTE

A proposta explorada neste trabalho propõe o QFC-LTE (*QoE Flow Control LTE*), capaz de realizar controle de fluxo e mapeamento de QoS entre a eNB e o UE baseado na busca de solução ótima da função objetivo $f_{QFC-LTE}(rv_n, ue_n, C_n, rb_n)$. O QFC-LTE é matematicamente formulado como em (3):

$$f_{QFC-LTE} = (rv_n \times ue_n) + [(1 - \lambda) \times (rb_n \times ue_n)] + C_n \quad (3)$$

onde rv_n representa a quantidade de recurso de Vídeo e VoIP destinado a uma conexão n , ue_n equivale ao número de UEs no sistema LTE destinado a n conexões, rb_n representa quantidade de recurso de BE (*Best Effort*) destinado a uma conexão n e C_n representa o nível de congestionamento na rede. O fator λ é uma variável que assume o valor 1 se a conexão é alocada somente para os tráfegos prioritários, e recebe 0, caso a conexão é alocada para todos os tráfegos.

Com base no conjunto de requisitos da função objetivo, o controle de fluxo é determinado e baseado na solução ótima esperada. Este processo é atingido quando se consiste em maximizar a função objetivo dada por (4):

$$\max f_{QFC-LTE}(rv_n, ue_n, C_n, rb_n) \quad (4)$$

Considerando este método de controle de fluxo, quando um novo fluxo é estabelecido ocorre um novo mapeamento de QoS entre a eNB e os UEs, otimizando desta forma as aplicações de Vídeo, VoIP e FTP e os fluxos são classificados e escalonados conforme a solução ótima do método proposto.

6. Simulações e Resultados

Para avaliar esse trabalho, foram realizadas simulações conduzidas no módulo LTE, no NS-2 (*Network Simulator 2*), versão 2.33, desenvolvido por [Qiu *et al* 2009]. Para transmissão de vídeos foi utilizado o *framework* Evalvid [Evalvid 2012]. O modelo de topologia simulado para a rede LTE está ilustrado na Figura 1.

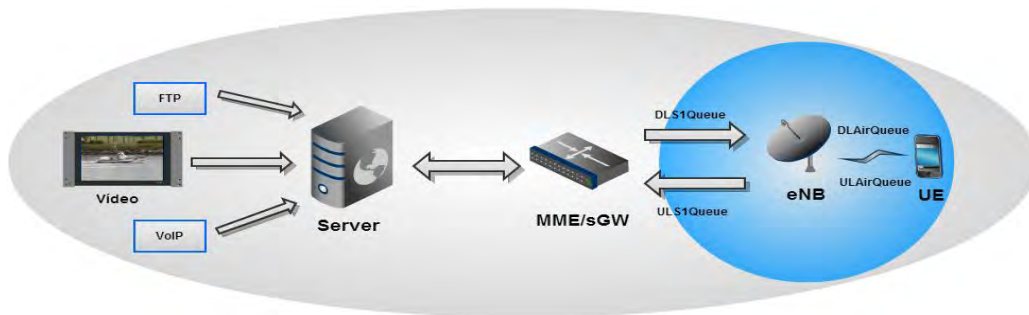


Figura 1. Cenário da Simulação para o sistema LTE.

Na simulação, o fluxo de vídeo utilizado foi o *Coastguard*, taxa de dados de 512 kbps, composto de 596 quadros, resolução de 352 x 288, 30 fps e GoP: 5 e 30. Além das especificações e padrões de vídeo, FTP e VoIP, é apresentado na Tabela 1 os principais parâmetros utilizados na simulação. Os parâmetros mostrados nesta tabela estão relacionados ao ambiente de simulação proposto em [Qiu *et al* 2010]. Além das especificações e padrões de vídeo, FTP e VoIP, a Tabela 1 apresenta os principais parâmetros utilizados na simulação. Os parâmetros mostrados nesta tabela estão relacionados ao ambiente de simulação proposto em [Qiu *et al* 2010].

Tabela 1. Parâmetros de simulação.

Parâmetros	Valor Utilizado
Vídeo	Taxa 512 kbps
Resolução	352 x 288
Total de Quadros	596 quadros
Taxa de Frames	30 fps
GoP	5 e 30
FTP	Taxa 512 kbps
VoIP	Taxa 8 kbps
Tempo de Simulação	19.9 s
Número de Simulações	100
Intervalo de Confiança	95%
Framework de Vídeo	Evalvid
Simulador	NS-2 2.33 e Módulo LTE

Para as métricas coletadas na simulação, utilizou-se um nível de confiança de 95% para análise dos resultados, além de analisar os dados quantitativos das métricas de QoE para PSNR, SSIM e VQM em sua média, desvio padrão e variância.

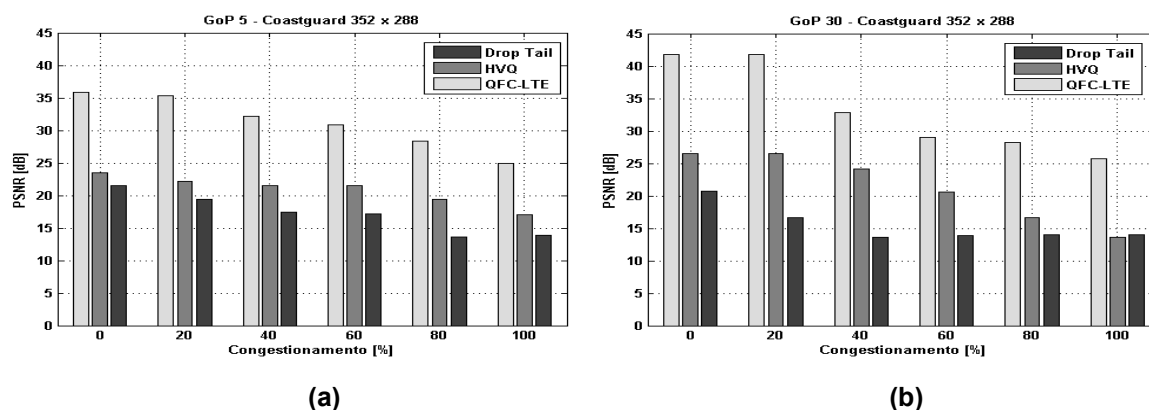


Figura 2. PSNR para fluxo de vídeo *Coastguard* 352 x 288: (a) GoP 5 e (b) GoP 30.

A primeira análise realizada foi quanto a métrica de PSNR, considerando diferentes níveis de congestionamento (0 – 100%) durante o processo de transmissão do vídeo *Coastguard* no sistema LTE, conforme é ilustrado na Figura 2 (a) GoP 5 e Figura 2 (b) GoP 30.

Os resultados demonstram que o controle de fluxo QFC-LTE, entre a eNB e o UE foi mais efetivo que o controle de fluxo *Drop Tail*, apresentando uma excelente qualidade de vídeo, para o QFC-LTE, segundo mapeamento de valores de PSNR para MOS (*Mean Option Score*), Tabela 2 [Piamrat *et al* 2009].

Tabela 2. Mapeamento do PSNR para o MOS.

PSNR (dB)	> 37	31 – 37	25 – 31	20 – 25	< 20
MOS	Excelente (5)	Bom (4)	Regular (3)	Pobre (2)	Péssimo (1)

Para o controle de fluxo proposto QFC-LTE, a qualidade do vídeo apresentou um nível de Bom à Regular para o GoP 5 e de Excelente à Regular para o GoP 30, devido ao mapeamento de QoS entre a eNB e o UE baseado na busca de solução ótima da função objetivo proposta.

Comparando os níveis de PSNR para o GoP 5, Figura 2 (a), o pior caso foi para controle de fluxo *Drop Tail*, a diferença máxima atingiu 52,17% em 80% de congestão, comparado ao melhor caso, QFC-LTE. Para o GoP 30, Figura 2 (b) o nível de PSNR para o QFC-LTE apresentou ganho no melhor caso, de 60,15% em 20 % de congestão comparado ao pior caso para o controle de fluxo *Drop Tail*.

Na segunda análise realizada para o controle de fluxo quanto ao nível da métrica SSIM. A Figura 3 (a) GoP 5 e Figura 3 (b) GoP 30 apresentaram as distribuições dos níveis da métrica SSIM em relação à variação do congestionamento, para os controles de fluxo QFC-LTE, HVQ e *Drop Tail*.

A Figura 4 (a) GoP 5 e Figura 4 (b) GoP 30 é ilustrado as distribuições dos níveis da métrica VQM, para as mesmas condições de congestionamento na rede LTE das métricas anteriores. Os resultados comprovam que o controle de fluxo entre a eNB e o UE, provido pelo QFC-LTE é mais eficaz que os controles HVQ e *Drop Tail*.

Comparando o melhor caso, QFC-LTE com o nível de VQM para o GoP 5, pior caso, *Drop Tail*, a diferença máxima foi de 58,10% sem congestionamento, conforme ilustra a Figura 4 (a). Nota-se que, a partir da congestão de 40%, para o GoP 5, o nível da qualidade de vídeo para a métrica VQM aumenta, comprovando os resultados das métricas anteriores para o controle de fluxo QFC-LTE.

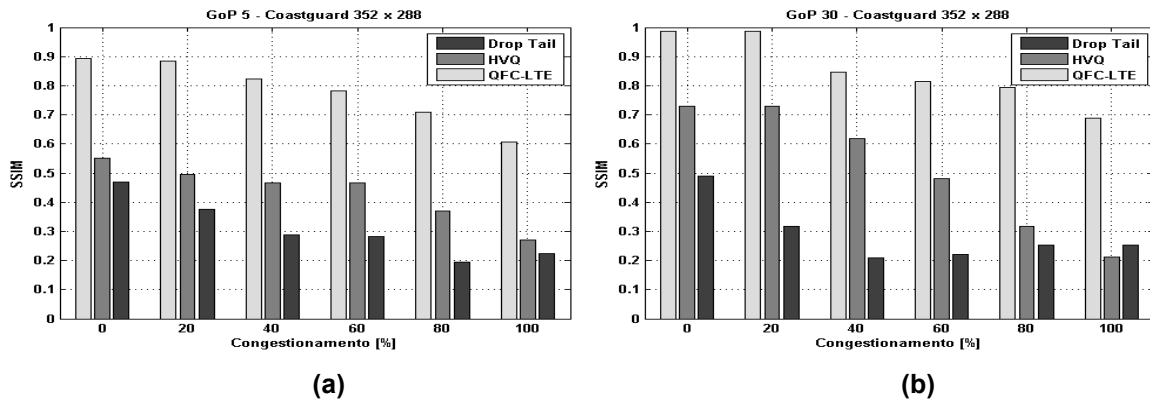


Figura 3. SSIM para fluxo de vídeo *Coastguard* 352 x 288: (a) GoP 5 e (b) GoP 30.

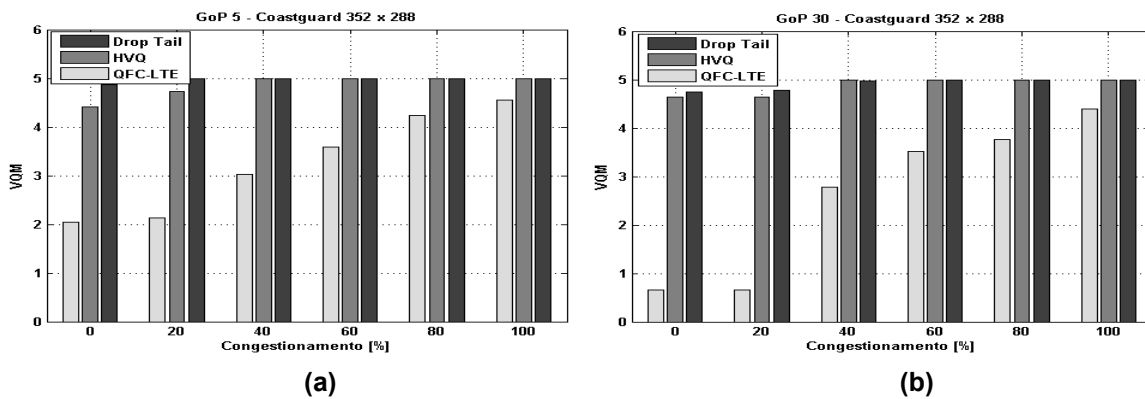
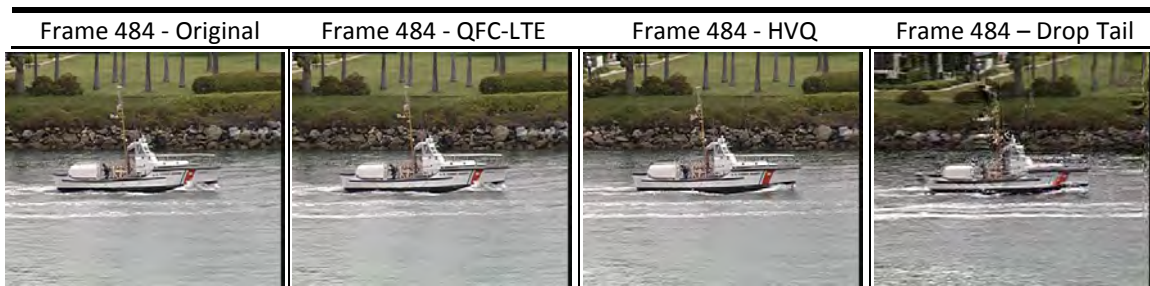


Figura 4. VQM para fluxo de vídeo *Coastguard* 352 x 288: (a) GoP 5 e (b) GoP 30.

Conforme a avaliação do vídeo e para validar a eficácia do controle de fluxo proposto, a Tabela 3 ilustra a sequência de quatro quadros de *frames* do vídeo *Coastguard*, *frame* 484, GoP 5 para 60% de congestionamento. Observa-se que, o vídeo recebido para o controle de fluxo QFC-LTE apresentou melhor qualidade de vídeo no fluxo entre eNB e o UE do sistema LTE.

Tabela 3. Frames do vídeo *Coastguard*, para o GoP 5 e congestão de 60%.

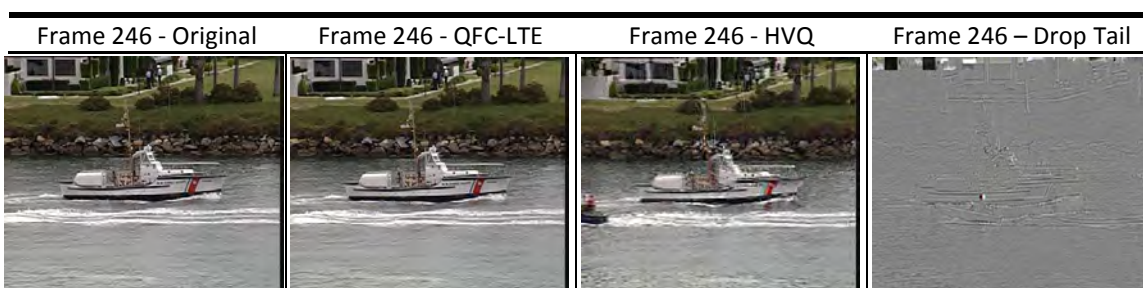


Os valores médios das métricas PSNR, SSIM e VQM para os *frames* da Tabela 3 são apresentados na Tabela 4, congestionamento de 60%, o vídeo recebido de acordo com o controle de fluxo no sistema LTE, para o QFC-LTE, HVQ e *Drop Tail*, GoP 5 apresentam, respectivamente, qualificação do vídeo como Bom, Pobre e Péssimo.

Tabela 4. Resultados das métricas de QoE para GoP 5, congestão em 60%.

	PSNR [dB]	SSIM	VQM	Qualidade
QFC-LTE	31	0,78	3,58	Bom
HVQ	21,45	0,47	5	Pobre
Drop Tail	17,22	0,28	5	Péssimo

A Tabela 5 ilustra a sequência de quatro quadros de *frames* do vídeo *Coastguard*, *frame* 246, GoP 30 para 60% de congestionamento. Verifica-se que os vídeos recebidos para o controle de fluxo *Drop Tail* e HVQ foi bastante degradado. O mesmo não acontece quando o sistema LTE utilizou o controle de fluxo QFC-LTE,

Tabela 5. Frames do vídeo *Coastguard*, para o GoP 30 e congestão de 60%.

Os valores médios das métricas PSNR, SSIM e VQM para os *frames* da Tabela 5 são apresentados na Tabela 6, congestionamento de 60% para o GoP 30, o vídeo recebido de acordo com o controle de fluxo do sistema LTE, para o QFC-LTE, HVQ e *Drop Tail*, apresentam, respectivamente, qualificação do vídeo como Regular, Pobre e Péssimo.

Tabela 6. Resultados das métricas de QoE para GoP 30, congestão em 60%.

	PSNR [dB]	SSIM	VQM	Qualidade
QFC-LTE	29	0,81	3,52	Regular
HVQ	20,64	0,48	5	Pobre
Drop Tail	13,85	0,28	5	Péssimo

Os resultados mostram que, a medida que utilizou-se o controle de fluxo QFC-LTE, a perda de *frames* foi bem menor, apresentando um impacto bem melhor na qualidade do vídeo. O controle de fluxo, aliado ao mapeamento de QoS entre o eNB e o UE, melhorou as métricas de QoE, para o GoP 5 e GoP 30. Isso devido a importância do agrupamento de quadros, além do processo de atualização mais rápida dos *frames*.

7. Conclusões e trabalhos futuros

Observamos que as métricas de QoE para a transmissão de conteúdo multimídia apresentaram melhores níveis para o controle de fluxo QFC-LTE e níveis mais baixos para os controles de fluxo HVQ e *Drop Tail*, devido a maior dependência de pacotes, codificação baixa, capacidade máxima esgotada da política de controle do *Drop Tail* (novos pacotes que chegam são descartados) e da falta de mapeamento de QoS no HVQ. A proposta QFC-LTE aumentou significativamente os níveis de QoE se comparado com os controles de fluxo HVQ e *Drop Tail* para o GoP 30, a porcentagem

máxima, entre o melhor caso, QFC-LTE e o pior caso, *Drop Tail* foi de 86,11% para a métrica VQM, na congestão de 20%.

Para trabalhos futuros, propõe-se um estudo utilizando outras métricas de QoE, entre elas, métricas subjetivas como, por exemplo, o MOS. Propõe-se também, a reformulação do método de controle de fluxo QFC-LTE, para considerar os diferentes valores dos GoPs para provisão de QoE no sistema LTE.

Referências

- Begazo, D., Rodriguez, D. e Ramirez, M. (2011) “Avaliação de Qualidade de Vídeo sobre uma Rede IP usando Métricas Objetivas”, In: Conferência Iberoamericana em Sistema, Cibernética e Informática CИСCI 2011, Orlando, USA, . v. I. p. 226-229.
- McQueen, D. (2009) “The momentum behind LTE adoption” IEEE Communications Magazine, vol. 47, no. 2, pp. 44–45, Feb.
- Rappaport, T. S. (2009) “Comunicações Sem Fio: Princípios e Práticas”, Pearson Prentice Hall, 2ª edição, São Paulo.
- Silva, D., Júnior, J., Coelho, M. e Dias, K. (2010) “Arquitetura Heterogênea com Gerenciamento da QoE e Suporte a Handover Transparente através de um Sistema Fuzzy-Genético”, In: XVI Simpósio Brasileiro de Sistemas Multimídias e Web – 2010, Belo Horizonte, MG.
- Hong, S. and Won, Y. (2010) “Incorporating Packet Semantics in Scheduling of Real-Time Multimedia Streaming”, Multimedia Tool Appl., vol. 46, pp. 463-492, Jan.
- Qiu, Q., Jian, C., Ping, L., and Pan, X. (2010) “Hierarchy Virtual Queue Based Flow Control in LTE/SAE”, Second International Conference on Future Networks ICFN’10, pp. 78 - 82.
- Holma, H. and Toskala, A. (2009) “LTE for UMTS OFDMA and SC-FDMA Based Radio Access”, John Wiley & Sons.
- Rodríguez, D. Z., and Bressan, G. (2012) “Video Quality Assessments on Digital TV and Video Streaming Services Using Objective Metrics”, IEEE Latin America Transactions, vol. 10, n. 1, Jan.
- Xinbo, G., Wen, L., Dacheng, T., and Xuelong, L. (2009) “Image Quality Assessment based on Multiscale Geometric Analysis”, Image Processing, IEEE Transactions on, IEEE, v.18, n.7, pp.1409-1423.
- Gelonch, A., Revés, X., Marojevic, V., Nasreddine, J., Pérez-Romero, J., and Sallent, O. (2010) “A Real Time Emulator Demonstrating Advanced Resource Management Solutions”, Wireless Personal Communications, Springer, v.54, n.1, pp. 123-136.
- Qiu, Q., Jian, C., Ping, L., Zhang, Q., and Pan, X. (2009) “LTE/SAE Model and its Implementation in NS 2”, Fifth International Conference on Mobile Ad-hoc and Sensor Networks MSN’ 09, pp. 299 - 303.
- Evalvid Tool (2012). Disponível em: www.tkn.tu-berlin.de/menue/research/evalvid.
- Piamrat, K., Viho, C., Ksentini, A., and Bonnin, J. (2009) “Quality of Experience Measurements for Video Streaming over Wireless Networks”, Sixth International Conference on Information Technology ITNG 09, pp. 1184 - 1189.

Aplicação de Processamento Paralelo em método iterativo para solução de sistemas lineares

Lauro C. M. de Paula¹, Leonardo B. S. de Souza², Leandro B. S. de Souza³,
Wellington S. Martins¹

¹Instituto de Informática – Universidade Federal de Goiás (UFG)
Caixa Postal 131 – 74.001-970 – Goiânia – GO – Brazil

²Escola de Engenharia Civil – Universidade Federal de Goiás (UFG)
Caixa Postal – 74.605-220 – Goiânia – GO – Brazil

³Departamento de Ciência da Computação – Universidade Federal de Minas Gerais
(UFMG) Caixa Postal – 31.270-901 – Belo Horizonte – MG – Brazil

{lauropaula, wellington}@inf.ufg.br, leobarra@eec.ufg.br, lbss@ufmg.br

Abstract. *This paper presents a C++ source code with parallel programming of the Hybrid Bi-Conjugate Gradient Stabilized (BiCGStab(2)) method, used for solving linear systems. The BiCGStab(2) method combines the Bi-Conjugate Gradient Stabilized (BiCGStab) method advantages, resulting in a convergence superior guarantee. The goal is to enable the rapid solution of linear systems to more complex problems (larger systems) can be solved in a short time. To validate the paper we used the parallel processing capacity of a GPU through the NVIDIA[®] Compute Unified Device Architecture (CUDA) and compared the sequential and parallelized code computational performance of the BiCGStab(2) applied to solving linear systems of varying sizes. There was a significant acceleration in tests with the parallelized code, which increases considerably as much as systems increase. It was possible to obtain performance gains for different systems size. The results allowed us to evaluate the developed code as a useful tool for solving large linear systems applied in various areas of science.*

Resumo. *Apresenta-se neste trabalho um código computacional em linguagem C++ com programação paralela do método Gradiente Bi-Conjugado Estabilizado Híbrido (BiCGStab(2)), utilizado para solução de sistemas lineares. O método em questão reúne as vantagens do método Gradiente Bi-Conjugado Estabilizado (BiCGStab), resultando em garantia superior de convergência. O objetivo é viabilizar a solução rápida de sistemas de equações lineares para que problemas cada vez mais complexos possam ser solucionados em um tempo curto. Para a validação do trabalho, utilizou-se a capacidade de processamento paralelo de uma GPU, por meio da arquitetura CUDA, e comparou-se o desempenho computacional dos códigos sequencial e paralelizado do BiCGStab(2) na solução de sistemas lineares de tamanhos variados. Observou-se uma aceleração significativa nos testes com o código paralelizado, que se acentua consideravelmente na medida em que os sistemas aumentam. Foi possível obter ganhos (speedup) de desempenho para diferentes tamanhos de sistema. Os resultados permitiram avaliar o código desenvolvido como uma ferramenta útil para solução de grandes sistemas lineares aplicados nas diversas áreas da ciência.*

1. Introdução

No campo de investigação da Dinâmica dos Fluidos Computacional (*Computational Fluid Dynamics* - CFD), existem diversos métodos utilizados para solução de sistemas lineares. Alguns deles são considerados ótimos em relação ao custo computacional, dependendo do tamanho do sistema a ser resolvido. Para trabalhos reais da ciência, nos quais os sistemas lineares a serem resolvidos podem ser muito grandes, o processamento computacional pode durar vários dias e as diferenças de velocidade de solução dos métodos são, definitivamente, significantes. Com isso, a implementação de métodos robustos e eficientes é importante e muitas vezes indispensável, para que as simulações de escoamentos de fluidos sejam realizadas com qualidade e em tempo não proibitivo.

Como os sistemas lineares encontrados na CFD são, normalmente, grandes e esparsos, métodos iterativos de solução são mais indicados do que métodos exatos, também por usarem menos memória e reduzirem erros de arredondamento do computador [Franco 2006]. Métodos iterativos realizam aproximações sucessivas, em cada iteração, para obter uma solução mais precisa para o sistema. Os Métodos iterativos clássicos como o de Jacobi e de Gauss-Seidel, apesar de sua fácil implementação, podem apresentar convergência lenta ou mesmo não convergirem para grandes sistemas [Versteeg and Malalasekera 1995]. Portanto, a pesquisa e a comparação de métodos iterativos a serem implementados em códigos computacionais podem ser consideradas tarefas importantes tanto na CFD quanto em outras áreas da ciência que envolvam a solução de grandes sistemas de equações lineares.

Trabalhos recentes fizeram uso de GPU (*Graphics Processing Unit*), por meio da arquitetura CUDA (*Compute Unified Device Architecture*), desenvolvida e mantida pela NVIDIA[®], para solucionar sistemas lineares. Elise Bowins [Bowins 2012], por exemplo, apresenta uma comparação de desempenho computacional entre o método de Jacobi e o método Gradiente Bi-Conjugado Estabilizado (BiCGStab), desenvolvido por Van der Vorst [Vorst 1992]. Ambos os métodos foram implementados em uma versão sequencial e outra paralelizada, e foi possível concluir que, conforme o tamanho do sistema aumenta, a implementação usando CUDA supera a implementação sequencial.

Nesse contexto, este trabalho apresenta um código computacional em linguagem C++ com programação paralela do método Gradiente Bi-Conjugado Estabilizado Híbrido (BiCGStab(2)). Tal método foi implementado em uma versão sequencial e outra paralelizada. Para a versão sequencial, foi utilizada a linguagem computacional C++, e para a paralelizada foi utilizada uma extensão da linguagem C, CUDA-C. O objetivo foi viabilizar a solução rápida de sistemas lineares para que problemas cada vez mais complexos (sistemas maiores) possam ser solucionados em um espaço curto de tempo. Para tal, foi utilizada a capacidade de processamento paralelo de uma unidade de processamento gráfico (GPU), por meio da arquitetura CUDA, e comparou-se o desempenho computacional dos códigos sequencial e paralelizado do BiCGStab(2) na solução de sistemas lineares de tamanhos variados. Tanto quanto sabemos, não se tem conhecimento de outros trabalhos publicados com programação paralela do método BiCGStab(2). Portanto, este artigo não faz comparação da versão paralelizada do BiCGStab(2) com implementações de outros trabalhos.

Este artigo está organizado da seguinte forma. Na Seção 2, é detalhado o método iterativo BiCGStab(2). Na Seção 3, descreve-se o material e os métodos utilizados para

se alcançar o objetivo do trabalho. Os resultados são discutidos na Seção 4. A Seção 5 traz as conclusões e algumas propostas para trabalhos futuros.

2. Método BiCGStab(2)

Resolver um sistema linear $\mathbf{Ax} = \mathbf{b}$ para o vetor $\mathbf{x}$, onde $\mathbf{A}$ é uma matriz de n linhas e n colunas e $\mathbf{b}$ um vetor de n linhas e 1 coluna, pode consumir bastante tempo, principalmente quando $\mathbf{A}$ é grande. Existem diversos métodos iterativos para resolver esse sistema linear. Tais métodos realizam aproximações sucessivas, em cada iteração, para obter uma solução mais precisa para o sistema [Bowins 2012]. Ainda, métodos iterativos são recomendados para sistemas lineares grandes com matrizes esparsas [Molnarka and Miletics 2005].

Os métodos iterativos podem ser classificados em dois grupos: os métodos estacionários e os métodos não-estacionários. Nos métodos estacionários, a mesma informação é usada em cada iteração. Como consequência, estes métodos são, normalmente, mais fáceis de implementar. Nos métodos não-estacionários, a informação usada pode mudar a cada iteração. Esses métodos são mais difíceis de implementar, mas, geralmente, fornecem uma convergência mais rápida para o sistema. Além disso, métodos iterativos se mostram mais adequados mesmo quando a matriz dos coeficientes é densa [Vorst 1995].

Neste trabalho, foca-se em um método iterativo não-estacionário, chamado método do Gradiente Bi-Conjugado Estabilizado Híbrido (BiCGStab(2)). O método BiCGStab(2) foi desenvolvido por Van der Vorst e Sleijpen [Sleijpen and Vorst 1995]. De acordo com seus criadores, tal método reúne as vantagens do BiCGStab e do GMRES (*Generalized Minimum Residual*), resultando, normalmente, em um método com garantia de convergência superior à do BiCGStab e adequado, por exemplo, para solução de sistemas lineares gerados na solução das equações diferenciais de escoamentos de fluidos.

1. $\mathbf{r}_0 = \mathbf{b} - \mathbf{Ax}_0$	20. $\mathbf{v} = \mathbf{s} - \beta\mathbf{v}$
2. Escolha um vetor arbitrário $\hat{\mathbf{r}}_0$	21. $\mathbf{w} = \mathbf{Av}$
tal que $(\mathbf{r}_0, \hat{\mathbf{r}}_0) \neq 0$, como $\hat{\mathbf{r}}_0 = \mathbf{r}_0$	22. $\gamma = (\mathbf{w}, \hat{\mathbf{r}}_0)$
3. $\rho = \alpha = \omega_1 = \omega_2 = 1$	23. $\alpha = \rho/\gamma$
4. $\mathbf{w} = \mathbf{v} = \mathbf{p} = \mathbf{0}$	24. $\mathbf{p} = \mathbf{r} - \beta\mathbf{p}$
5. For $i = 0, 2, 4, 6, \dots$	25. $\mathbf{r} = \mathbf{r} - \alpha\mathbf{v}$
6. $\rho' = -\omega_2\rho$	26. $\mathbf{s} = \mathbf{s} - \alpha\mathbf{w}$
Even Bi-CG step	27. $\mathbf{t} = \mathbf{As}$
7. $\rho = (\mathbf{r}_i, \hat{\mathbf{r}}_0)$	GMRES(2)-part
8. $\beta = \alpha\rho/\rho'$	28. $\omega_1 = (\mathbf{r}, \mathbf{s})$
9. $\rho' = \rho$	29. $\mu = (\mathbf{s}, \mathbf{s})$
10. $\mathbf{p} = \mathbf{r}_i - \beta(\mathbf{p} - \omega_1\mathbf{v} - \omega_2\mathbf{w})$	30. $\mathbf{v} = (\mathbf{s}, \mathbf{t})$
11. $\mathbf{v} = \mathbf{Ap}$	31. $\tau = (\mathbf{t}, \mathbf{t})$
12. $\gamma = (\mathbf{v}, \hat{\mathbf{r}}_0)$	32. $\omega_2 = (\mathbf{r}, \mathbf{t})$
13. $\alpha = \rho/\gamma$	33. $\tau = \tau - v^2/\mu$
14. $\mathbf{r} = \mathbf{r}_i - \alpha\mathbf{v}$	34. $\omega_2 = (\omega_2 - v \cdot \omega_1/\mu)/\tau$
15. $\mathbf{s} = \mathbf{Ar}$	35. $\omega_1 = (\omega_1 - v \cdot \omega_2)/\mu$
16. $\mathbf{x} = \mathbf{x}_i + \alpha\mathbf{p}$	36. $\mathbf{x}_{i+2} = \mathbf{x} + \alpha\mathbf{p} + \omega_1\mathbf{r} + \omega_2\mathbf{s}$
Odd Bi-CG step	37. $\mathbf{r}_{i+2} = \mathbf{r} - \omega_1\mathbf{s} - \omega_2\mathbf{t}$
17. $\rho = (\mathbf{s}, \hat{\mathbf{r}}_0)$	38. Se $\mathbf{x}_{i+2}$ for suficientemente
18. $\beta = \alpha\rho/\rho'$	preciso: parar.
19. $\rho' = \rho$	39. End For

Figura 1. Algoritmo do método iterativo BiCGStab(2), adaptado de [Sleijpen and Vorst 1995].

É representado na Figura 1 o algoritmo do método BiCGStab(2). Algumas adaptações de nomenclatura foram feitas com relação ao algoritmo original, na busca de melhor entendimento do método. No algoritmo, as letras gregas representam escalares, letras minúsculas representam vetores expressos na forma matricial, as letras maiúsculas representam matrizes e os parênteses com vetores separados por vírgula representam produtos escalares entre os vetores.

No passo 38 do método, para que o vetor x_{i+2} seja suficientemente preciso, o maior valor referente à diferença entre os resultados de cada termo do vetor x em duas iterações consecutivas, dividida pelo resultado do termo na iteração atual deverá ser menor do que uma dada precisão, fornecida inicialmente como, por exemplo, 10^{-5} . Ou seja, $\max(\frac{x_i - (x_{i-1})}{x_i}) < 0.00001$.

3. Material e Métodos

A Unidade de Processamento Gráfico de Propósito Geral (ou GPGPU) utiliza a GPU não apenas para renderização gráfica, mas também para processamento de imagem, visão computacional, cálculo numérico, dentre outras aplicações. A GPU foi inicialmente desenvolvida como uma tecnologia orientada à vazão, otimizada para cálculos de uso intensivo de dados, onde muitas operações idênticas podem ser realizadas em paralelo sobre diferentes dados.

Diferentemente de uma CPU (*Central Processing Unit*) Multicore, a qual executa, normalmente, algumas *threads* em paralelo, a GPU foi projetada para executar milhares de *threads* em paralelo. Modelos de programação, tal como CUDA [CUDATM 2011] e OpenCL [Tsuchiyama et al. 2010], permitem que as aplicações sejam executadas mais facilmente na GPU [Xiao and Feng 2009]. CUDA foi a primeira arquitetura e interface para programação de aplicação (*Application Programming Interface* - API), criada pela NVIDIA[®] em 2006, a permitir que a GPU pudesse ser usada para uma ampla variedade de aplicações [Bowins 2012]. Em CUDA, a CPU é chamada *host*, onde toda a parte sequencial do código é executada e as funções *kernel* são chamadas para serem executadas na GPU. A GPU, por sua vez, é chamada *device*, a qual pode executar todas as partes aritmeticamente intensivas do código [Gravvanis and Papadopoulos 2011]. Neste trabalho, o algoritmo do método BiCGStab(2), descrito na seção 2, foi codificado em linguagem C++ empregando uma implementação inteiramente sequencial, bem como uma implementação parcialmente paralelizada empregando CUDA-C [CUDATM 2009].

Ao analisar o algoritmo do método BiCGStab(2), é possível verificar uma forte dependência de dados, isto é, alguns passos não podem ser executados antes que outros sejam executados anteriormente. Devido a essa característica, apenas algumas partes do método foram paralelizadas, como: multiplicação entre matriz e vetor, soma de vetores, subtração de vetores e multiplicação de vetor por escalar. Por exemplo, o passo 36 do algoritmo do BiCGStab(2) é dividido entre n *threads*, onde n é o tamanho do vetor x e para cada $0 \leq i < n$, uma *thread* executa a seguinte operação:

$$x_i = x_i + \alpha p_i + \omega_1 r_i + \omega_2 s_i. \quad (1)$$

Portanto, cada elemento do vetor x é calculado em paralelo. O passo 11 ($v = Ap$) somente poderá ser executado após o passo 10 ter sido totalmente executado. Para

garantir essa condição, utiliza-se uma barreira de sincronização, por meio de uma função padrão (*cudaDeviceSynchronize()*) da arquitetura CUDA. De forma análoga, o mesmo ocorre para outras etapas do algoritmo.

Os produtos escalares são realizados sequencialmente. Poderia ser possível aumentar o paralelismo dividindo os produtos entre múltiplas somas feitas em paralelo, mas isso pode exigir muita sobrecarga (*overhead*) para resultar em um ganho de desempenho significativo [Bowins 2012].

A seguir, mostra-se um pequeno trecho do código em CUDA-C usado para a implementação do BiCGStab(2). Inicialmente, o código é executado na CPU, onde as funções *kernel* são chamadas para serem executadas na GPU. A multiplicação entre matriz (A) e vetor (x), por exemplo, acontece em duas etapas, por meio de duas chamadas consecutivas à duas funções *kernel* distintas. Na primeira função *kernel*, uma matriz auxiliar (C) recebe a multiplicação de cada elemento da matriz pela respectiva posição do vetor. Para isso, n threads executam em paralelo, onde cada thread percorre uma linha da matriz. Na segunda função, um vetor auxiliar (aux) recebe a soma dos elementos de cada linha da matriz C , obtendo o vetor resultante da multiplicação.

multiplicacao1 <<< $nBlocos, nThreadsPorBloco$ >>> (A, x, C, n); (2)

multiplicacao2 <<< $nBlocos, nThreadsPorBloco$ >>> (C, aux, n); (3)

Para cada uma das funções *kernel*, $nBlocos = nThreadsPorBloco = \sqrt{n}$. Ou seja, o número de blocos e o número de threads por bloco, especificados na chamada da função, são iguais a raiz quadrada do número de linhas (n) da matriz A . Se $\sqrt{n}$ não é um número inteiro, então esse valor é arredondado para cima.

Existem alguns passos separados do método BiCGStab(2) que poderiam ser executados em paralelo, porém isso não é possível utilizando uma única GPU. Uma possível solução seria utilizar múltiplas GPUs ou, provavelmente, uma próxima geração de GPUs. Entretanto, a maioria dos passos do algoritmo devem ser realizados em sequência.

Os resultados obtidos com o método BiCGStab(2) paralelizado foram confrontados com os do BiCGStab(2) sequencial, com o intuito de verificar o ganho computacional obtido com a implementação paralelizada. Para avaliar o ganho computacional obtido por meio da implementação em CUDA, registrou-se o tempo dispendido em cada iteração do algoritmo. Todos os testes foram realizados em um computador com processador Intel Core i7 2600 3,4GHz, 8 GB de memória RAM, com placa de vídeo NVIDIA® GeForce GTX 550Ti dispondo de 2 GB de memória e 192 processadores operando a 900MHz.

4. Resultados

Todos os sistemas lineares utilizados neste trabalho foram gerados através de funções do software MATLAB (versão R2011a). A matriz (A) dos coeficientes de cada sistema foi gerada de forma aleatória utilizando a função padrão *gallery('dorr', n)*, que retorna uma matriz quadrada, de dimensão n , esparsa e diagonal dominante. O vetor (x) das incógnitas foi gerado de forma aleatória por meio da função padrão *randn(n, 1)*, que retorna um vetor de n linhas e 1 coluna. O vetor (b) dos termos independentes foi gerado multiplicando-se a matriz A e o vetor x ($b = A * x$). Então, para cada um dos sistemas gerados, foi

repassado para o método BiCGStab(2), sequencial e paralelizado, apenas a matriz A e o vetor b , que, após a tentativa de convergência do sistema, retornou o vetor x .

Os gráficos comparativos do tempo de processamento (em segundos) dos diversos sistemas lineares resolvidos com o método BiCGStab(2) são apresentados nas figuras 2 e 3.

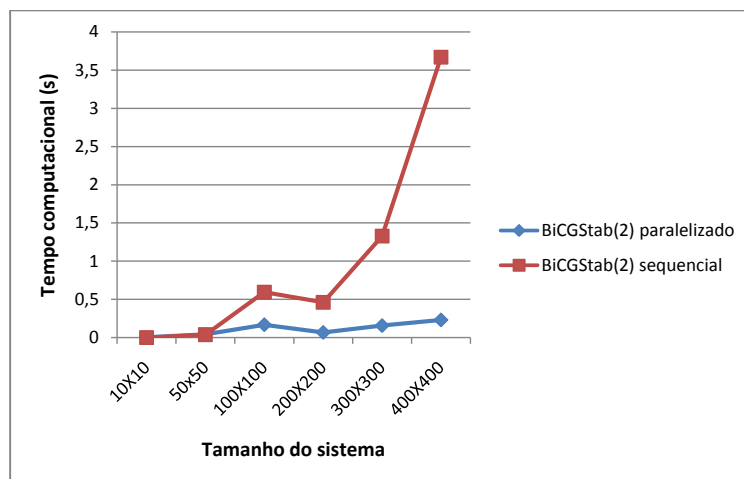


Figura 2. Comparação da velocidade de cálculo para os sistemas lineares de dimensão entre 10x10 e 400x400.

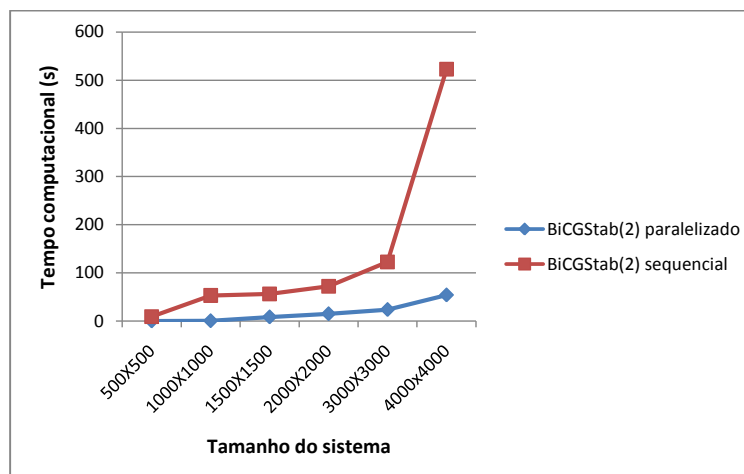


Figura 3. Comparação da velocidade de cálculo para os sistemas lineares de dimensão entre 500x500 e 4000x4000.

4.1. Avaliação do ganho computacional proporcionado pela versão paralelizada do BiCGStab(2)

Analisando os gráficos nas figuras 2 e 3, nota-se a superioridade do método BiCGStab(2) paralelizado na solução dos sistemas tratados. É possível verificar que a versão sequencial do método requer um esforço computacional que aumenta rapidamente com o tamanho

do sistema. Observa-se também que o tempo para a versão paralelizada também aumenta nesse caso, porém com taxa de crescimento menos expressiva.

Para sistemas lineares com dimensão menor que aproximadamente 50×50 , a versão sequencial poderá se mostrar equivalente ou mais rápida que a paralelizada. Isso ocorre devido à existência de um *overhead* associado à paralelização das tarefas na GPU [Filho et al. 2011]. Ou seja, o tamanho do sistema a ser solucionado na GPU deve ser levado em consideração. Fatores em relação ao tempo de acesso à memória podem influenciar no desempenho computacional. Por exemplo, o acesso à memória global da GPU, geralmente, apresenta uma alta latência e está sujeito a um acesso aglutinado aos dados em memória [CUDATM 2011]. Entretanto, observa-se uma aceleração significativa nos testes com o BiCGStab(2) paralelizado, que se acentua consideravelmente na medida em que os sistemas aumentam.

5. Conclusões

Foi implementado e utilizado neste trabalho um código computacional do algoritmo do método iterativo BiCGStab(2), para solução de sistemas lineares de tamanhos variados. Tal método foi implementado em duas versões, uma sequencial e outra paralelizada, utilizando as linguagens C++ e CUDA-C, respectivamente. Na versão paralelizada, explorou-se o uso de uma GPU, por meio da arquitetura CUDA, para a paralelização de algumas etapas do algoritmo do BiCGStab(2).

Para os sistemas lineares avaliados neste artigo, constatou-se uma superioridade da versão paralelizada em relação ao tempo computacional gasto no cálculo de cada sistema. Foi possível obter ganhos (*speedup*) de desempenho, com a implementação usando CUDA, para diferentes tamanhos de sistema. Observou-se que a média de *speedup* ficou em torno de ≈ 10 . Portanto, concluiu-se que a versão paralelizada do BiCGStab(2) seria uma implementação mais apropriada, do ponto de vista computacional, desde que o tamanho do sistema empregado seja suficientemente grande para justificar o *overhead* gasto na paralelização das tarefas na GPU.

Trabalhos seguintes nessa linha de pesquisa poderão envolver solução de sistemas lineares ainda maiores, como, por exemplo, os sistemas gerados nas simulações de escoamentos de fluidos, amplamente estudados na Dinâmica dos Fluidos Computacional. Além disso, poderá ser possível aplicar uma otimização no código computacional da versão paralelizada do método BiCGStab(2), de tal forma que etapas que ainda são resolvidas sequencialmente poderão ser solucionadas em paralelo. Uma possível otimização seria a utilização de *Persistent Threads* (PT), onde uma única função *kernel* é executada na GPU durante todo o processamento do algoritmo. Espera-se que nesse caso os ganhos computacionais proporcionados sejam ainda mais expressivos, devido à uma melhor utilização dos recursos de processamento paralelo da GPU. Adicionalmente, alternativas à arquitetura CUDA poderão ser investigadas para realização de estudos comparativos.

Agradecimentos

Os autores agradecem à CAPES pelo apoio fornecido à pesquisa.

Referências

- Bowins, E. C. (2012). A comparison of sequential and gpu implementations of iterative methods to compute reachability probabilities. *Proceedings First Workshop on GRAPH Inspection and Traversal Engineering*, pages 20–34.
- CUDATM, N. (2009). *NVIDIA CUDA C Programming Best Practices Guide*. NVIDIA Corporation, 2701 San Tomas Expressway Santa Clara, CA 95050.
- CUDATM, N. (2011). *NVIDIA CUDA C Programming Guide*. NVIDIA Corporation, 2701 San Tomas Expressway Santa Clara, CA 95050, 4.0 edition.
- Filho, A. R. G., Arruda, F. D. B., Harrop, R. K., and Yoneyama, T. (2011). Programacao paralela cuda para simulacao de modelos epidemiologicos baseados em individuos. *Simposio Brasileiro de Automacao Inteligente*.
- Franco, N. B. (2006). *Calculo Numerico*. Sao Paulo: Pearson Prentice Hall.
- Gravvanis, G. A. and Papadopoulos, C. K. F. (2011). Solving finite difference linear systems on gpus: Cuda based parallel explicit preconditioned biconjugate conjugate gradient type methods. *Journal of Supercomputing*, 61:590–604.
- Molnarka, G. and Miletics, E. (2005). A genetic algorithm for solving general system of equations. *Third Slovakian-Hungarian Joint Symposium on Applied Machine Intelligence*.
- Sleijpen, G. L. G. and Vorst, H. A. V. (1995). Hybrid bi-conjugate gradient methods for cfd problems. *Computational Fluid Dynamics REVIEW*, (902).
- Tsuchiyama, R., Nakamura, T., Iizuka, T., Asahara, A., Son, J., and Miki, S. (2010). *The OpenCL Programming Book*. Fixstars.
- Versteeg, H. K. and Malalasekera, W. (1995). *An introduction to computational fluid dynamics: the finite volume method*. Londres: Longman Scientific & Technical.
- Vorst, H. A. V. (1992). Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM Journal of Scientific and Statistical Computing*, 13(2):631–644.
- Vorst, H. A. V. (1995). Parallel iterative solution methods for linear systems arising from discretized pdes. *France Workshop Lecture Notes*.
- Xiao, S. and Feng, C. (2009). Inter-block gpu communication via fast barrier synchronization. *23rd IEEE International Parallel & Distributed Processing Symposium, ISSN: 15302075*.

Em Direção a uma Ferramenta de Suporte a Revisão Sistemática de Literatura

César Augusto Verçosa Mariano¹, Ivan do Carmo Machado², Luanna Lopes Lobato¹

¹Departamento de Ciência da Computação – Universidade Federal de Goiás
CEP: 75704-020 – Catalão – GO – Brasil

²Departamento de Ciência da Computação – Universidade Federal da Bahia
CEP: 40170-110 – Salvador – BA – Brasil

cesar.m@catalao.ufg.br, ivanmachado@dcc.ufba.br, lll@catalao.ufg.br

Abstract. *This article introduces the concept and the characteristics of Systematic Literature Review, explaining its advantages and difficulties for the researcher. Based on the understanding of Systematic Review, we propose a tool that aims to assist in this process. We illustrate the tool workflow with a running example, and discuss the expected effectiveness and efficiency gains regarding time to search for primary studies, and quality outcomes.*

Resumo. *Este artigo apresenta o conceito e as principais características da Revisão Sistemática de Literatura, explicitando suas vantagens e dificuldades para o pesquisador. A partir do entendimento de Revisão Sistemática, propomos uma ferramenta que busca auxiliar neste processo. Ilustramos o workflow da ferramenta com um exemplo prático, e discutimos os ganhos de eficiência e eficácia esperados para a ferramenta, em relação ao tempo de busca por estudos primários e a qualidade dos resultados.*

1. Introdução

O processo de revisão sistemática (RS) de artigos disponibilizados na literatura – do termo original Systematic Literature Review – teve origem em pesquisas da área médica, em que traçou-se como objetivo a elaboração de uma seleção completa e imparcial sobre tópicos de pesquisa, seguindo processos conhecidos e bem definidos, que propiciasse a sua replicabilidade (Pai, 2004). Tal processo tem sido adaptado para a área de Ciência da Computação, e nos últimos anos, é visível a sua utilização abrangente, especialmente na pesquisa em Engenharia de Software (Kitchenham, 2007).

Kitchenham (2007) define revisão sistemática de literatura como a identificação, validação e interpretação de toda pesquisa relevante (estudos primários) disponível para uma questão de pesquisa particular, ou um tópico de uma área, ou ainda um fenômeno de interesse. A principal razão para a realização da revisão sistemática é que esta possui uma estratégia de busca bem definida, o que contribui para a integridade das informações resultantes da pesquisa. Por sua natureza de completude e imparcialidade, pode-se considerar que a revisão sistemática agrega maior valor científico quando comparado a revisões executadas de forma *ad-hoc*. As principais vantagens do uso da revisão sistemática estão relacionadas à segurança no processo seguido, a possibilidade de replicabilidade do estudo, bem como, a confiabilidade do processo, uma vez que este é executado de modo planejado e sistemático.

No entanto, apesar das vantagens apresentadas, o processo manual de revisão sistemática é trabalhoso e propenso a falhas, caso não seja cuidadosamente seguido. Neste contexto, a presente investigação propõe o desenvolvimento de uma ferramenta para automatizar o processo de seleção e extração de dados de estudos primários, objetos de avaliação na revisão sistemática.

A ferramenta, intitulada como SysReTool (acrônimo para **S**ystematic **R**eview **T**ool), tem o objetivo de auxiliar o processo de extração dos dados, de modo a facilitar o gerenciamento das informações que são coletadas de artigos e, desta forma, otimizar o tempo de extração e análise, evitando a perda de informações relevantes.

O artigo está estruturado como segue. O processo de revisão sistemática é descrito na seção seguinte. A seção 3 apresenta a proposta da ferramenta. A seção 4 apresenta o cenário de execução da ferramenta. A seção 5 apresenta as considerações finais desta investigação.

2. Revisão Sistemática

Ao conduzir pesquisas, é necessário entender a área sob investigação, de forma a propor avanços que tenham alguma significância. Assim, informação oriunda de estudos primários deve ser coletada. Entretanto, à medida que a quantidade de estudos a serem analisados escala, a extração de dados tende a ser cansativa, o que leva a uma possível perda de dados relevantes.

Walia e Carver (2009) descrevem RS como *“um processo formalizado e repetitivo no qual os pesquisadores procuram sistematicamente um corpo literário para documentar o estado do conhecimento sobre um determinado assunto”*. Dentre os benefícios de se realizar uma RS, podemos considerar que ela fornece ao pesquisador maior confiança na revisão de informações relevantes, uma vez que a seleção dos estudos é feita de forma criteriosa em relação ao escopo da pesquisa.

Segundo Sampaio (2007), a RS é metódica, explícita e passível de reprodução. Uma RS requer perguntas claras em relação ao que se procura pesquisar, buscando ter respostas precisas. Como resultado, espera-se a identificação de dados relevantes ao domínio de pesquisa, extraídos através das respostas às perguntas definidas.

A RS tem como princípios gerais a exaustão na busca dos estudos analisados, a seleção justificada dos estudos por critérios de inclusão e exclusão explícitos e a avaliação da qualidade metodológica, bem como a quantificação do efeito dos tratamentos por meio de técnicas estatísticas (Lima, 2000).

De acordo com Kitchenham (2007), a execução de uma Revisão Sistemática engloba três fases principais: (i) *planejamento*, (ii) *realização*, e (iii) *relato* da revisão.

Os estágios associados com o planejamento da revisão são:

- **Identificação da necessidade de uma revisão:** os pesquisadores devem identificar e analisar revisões sistemáticas já existentes com base no fenômeno de interesse em relação aos critérios de avaliação adequados. A RS só deve ser feita se realmente houver necessidade.
- **Especificação das questões de pesquisa:** elas são o elemento principal de condução da metodologia de RS. O processo de pesquisa deve identificar estudos primários que tratam das questões de pesquisa. O processo de extração de dados deve extrair os itens de dados necessários para responder às

perguntas. O processo de análise de dados deve sintetizar os dados, de forma que as perguntas possam ser respondidas.

- **Desenvolvimento de um protocolo de revisão:** documento que especifica os métodos que serão utilizados para realizar a revisão, a fim de reduzir a possibilidade de parcialidade gerada pelas expectativas do investigador.
- **Avaliação do protocolo de revisão:** A consistência interna do protocolo pode ser verificada para confirmar que cada processo da pesquisa é derivado das questões de pesquisa; Os dados a serem extraídos irão resolver adequadamente as questões de pesquisa; O procedimento de análise dos dados é apropriado para responder às questões de pesquisa.

Após a definição do protocolo, os seguintes passos devem ser executados:

- **Seleção de estudos primários:** os critérios de seleção identificam os estudos primários que forneçam evidência direta sobre as questões de pesquisa. Uma vez que os estudos potencialmente relevantes primários foram obtidos, eles precisam ser avaliados quanto a sua relevância atual.
- **Avaliação da qualidade do estudo:** a avaliação da qualidade dos estudos primários proporciona critérios de inclusão e exclusão mais detalhados.
- **Extração e monitoramento de dados:** o objetivo desta etapa é criar formas de extração de dados para registrar com precisão as informações que os pesquisadores obtiveram dos estudos primários.
- **Síntese dos dados:** informações extraídas sobre os estudos (ex: objetivo, proposta, contexto de investigação, avaliação da qualidade do estudo, tamanhos de amostra, resultados, etc.) devem ser tabuladas de forma consistente com a questão de revisão. As tabelas devem ser estruturadas para destacar semelhanças e diferenças entre os resultados do estudo.

A fase final da revisão está associada com o relato da revisão:

- **Especificação de mecanismos de disseminação:** geralmente os resultados de uma revisão sistemática podem ser divulgados em revistas acadêmicas e/ou conferências. No entanto, pode haver outras formas de difusão do conhecimento resultante da investigação.
- **Avaliação do relatório:** especialistas podem avaliar o relatório final gerado.

É possível perceber que o processo de RS é dispendioso e propenso a falhas, principalmente quando executado manualmente e por pessoas não especialistas. Por isso, neste trabalho é proposto o desenvolvimento de uma ferramenta para auxiliar este processo, como descrito na próxima seção.

3. Ferramenta proposta: SysReTool

Baseado nos estudos analisados (Kitchenham, 2007) (Walia, 2009) (Lobato, 2012), pôde-se observar alguns passos que devem ser executados para que haja uma eficiente e eficaz análise sistemática da literatura. Neste sentido, esses passos devem ser refletidos como funcionalidades da ferramenta proposta. Lobato (2012) apresenta um conjunto de etapas fundamentais para a condução de uma RS, são eles:

A 1ª Etapa consiste na identificação das questões de pesquisa. Ela engloba a definição das palavras-chave a serem utilizadas na busca de estudos primários. Este é o ponto de partida do processo, uma vez que é necessário que estabelecer um propósito da

pesquisa bem definido, de forma que se consiga definir questões que resultem em dados relevantes para a pesquisa. A 2ª Etapa seleciona as fontes de pesquisa que serão utilizadas para busca dos estudos primários.

A 3ª Etapa envolve a identificação dos estudos primários. É importante definir os critérios de inclusão e exclusão dos estudos, possibilitando assim que os estudos mais relevantes ao tema de pesquisa sejam incluídos. A seleção consiste na análise criteriosa de diferentes seções do artigo, de forma a elencar os estudos primários candidatos, como segue:

- **Selecionar os estudos pelo título:** os estudos obtidos em anais de conferências, periódicos, e nos motores de busca¹ são selecionados de acordo com o título (ou palavras chave). Se o título do estudo não condiz com o título, ou palavras chave fornecidas pelo usuário, não será incluído na lista de estudos armazenados pela ferramenta.
- **Eliminar trabalhos duplicados:** os estudos já armazenados pela ferramenta são verificados. Se há estudos com o mesmo nome, ou nomes bastante parecidos, um deles é excluído. Se dois trabalhos diferentes possuem a mesma informação o mais completo é considerado.
- **Selecionar estudos por critérios definidos:** o usuário pode filtrar os resultados obtidos na etapa anterior. Os estudos podem ser visualizados pela fonte em que foram identificados (motores de busca, conferências, periódicos) pelo ano (e/ou intervalo) de publicação, ou pelo nome do autor.
- **Selecionar trabalhos por resumo:** o usuário lê o resumo dos estudos restantes após as etapas anteriores e exclui, manualmente, os que ele julgar não serem relevantes para seu estudo.

A 4ª Etapa consiste na leitura e análise do artigo. O revisor faz a leitura de todo o texto e exclui, manualmente, os que ele julgar não serem relevantes para seu estudo. Cada estudo é avaliado de acordo com os critérios de qualidade. Os critérios de qualidade são necessários para pontuar os estudos selecionados, de modo a definir um ranking de relevância entre os estudos. Os critérios devem ser definidos com base nas questões de pesquisa estabelecidas na 1ª Etapa. Os estudos são pontuados de acordo com uma escala de relevância: *baixa*, *média* e *alta*. Para facilitar o gerenciamento de acordo com a pontuação, pode-se atribuir valores, e.g., *0(zero)* para baixa relevância, *0,5* para média e *1* para alta (Lobato, 2012).

A 5ª Etapa inclui a extração dos dados provenientes dos estudos selecionados. Assim, as informações relevantes são recolhidas através da análise dos trabalhos, a fim de responder às questões de pesquisa. As respostas às questões devem ser descritivas o suficiente para que, ao lê-las, o usuário consiga elaborar estratégias de solução e identifique falhas em relação ao campo de estudo.

A 6ª Etapa apresenta a síntese da pesquisa, que envolve a identificação dos resultados que podem ser utilizados como análise, sendo ressaltadas as vantagens e desvantagens em relação a área sob investigação, analisadas durante a revisão.

¹Os motores de busca comumente utilizados para a recuperação de publicações em Engenharia de Software são: ACM, IEEEEXPLORE, SpringerLink, ScienceDirect, e Scopus. Eventualmente, considera-se Google Scholar como um motor relevante, embora ele indexe os artigos listados pelos motores acima, o que pode resultar em redundância.

A ferramenta SysReTool apresenta dois estágios de execução: automático e manual. A execução automatizada tem por objetivo tornar a pesquisa mais otimizada, evitando que o usuário execute todas as tarefas; na manual, o usuário deve verificar cada parte da análise sem o intermédio de atividades automatizadas, apenas tendo a ferramenta como suporte ao gerenciamento das tarefas.

Inicialmente a ferramenta SysReTool executará uma busca automática nos motores de busca, e nos anais de conferências e periódicos— conforme definidos na 2ª Etapa. De modo adicional, o usuário poderá digitar links de sites onde deseja fazer a busca. Tais funcionalidades são ilustradas na Figura 1, que apresenta uma das telas do protótipo da ferramenta.

Figura 1. Tela do protótipo da ferramenta – funcionalidade que demonstra como o usuário inicia a pesquisa.

Os estudos, encontrados a partir das fontes sugeridas e das adicionais, serão filtrados por títulos de acordo com as palavras chave que foram informados pelo usuário na 1ª Etapa. Após este primeiro passo, é apresentada uma lista com os estudos selecionados de forma automatizada, mostrando o título, o ano e o autor. Ao clicar no título é possível visualizar o artigo por completo. Assim, o usuário tem a opção de remover, de forma manual, os artigos selecionados que não são interessantes para compor o conjunto de estudos que serão utilizados na pesquisa.

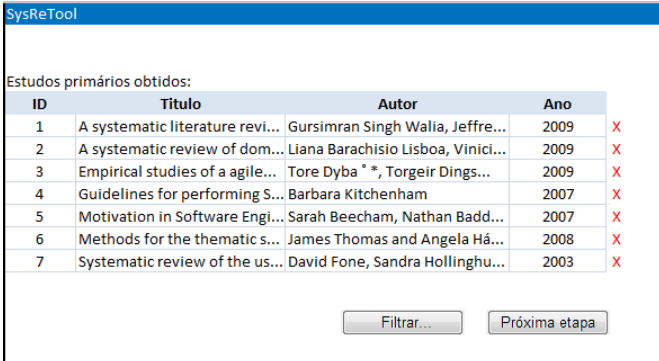
É importante ressaltar que durante a identificação dos estudos de forma automatizada, o trabalho do usuário foi reduzido e o tempo que o usuário gastaria lendo títulos foi poupado, o que poderia ocasionar cansaço, resultando na possível exclusão indevida de estudos que poderiam ser relevantes.

Após a busca automática dos estudos, a ferramenta SysReTool apresentará todos os estudos selecionados, conforme ilustrado na Figura 2, fornecendo a opção de que os estudos repetidos sejam removidos do conjunto de estudos candidatos. Esta etapa é automatizada, auxiliada pela ferramenta. O revisor pode ainda visualizar os estudos e, caso algum estudo não o interesse, é possível executar a exclusão de modo manual, clicando no X, ao lado das informações sobre o estudo.

Adicionalmente, a SysReTool possibilita que o usuário exclua artigos de acordo com o número de páginas. Essa funcionalidade, que está disponível ao clicar no botão *Filtrar*, é importante já que é possível desconsiderar estudos com número reduzido de páginas, por não estarem completos o suficiente para responder às questões de pesquisa.

Este tipo de seleção, que possibilita a exclusão dos estudos que não atendem a especificação definida de ano, autor, etc., é feita em nível da solicitação do usuário. Ou

seja, os estudos que estão fora do *range* de seleção não serão completamente excluídos, uma vez que estes são relevantes ao campo de pesquisa de um modo geral. Os estudos ficam armazenados para que a ferramenta possa desfazer a ação do usuário, caso ele deseje restaurar alguma entrada que tenha sido excluída, ou deseje fazer a seleção dos estudos considerando outras opções. Assim, para cada seleção, de acordo com a opção escolhida, os estudos serão apresentados e a lista dessa seleção pode ser salva para futuras análises, ou ainda para modificar a definição das opções de seleção.



ID	Título	Autor	Ano
1	A systematic literature revi...	Gursimran Singh Walia, Jeffre...	2009
2	A systematic review of dom...	Liana Barachisio Lisboa, Vinici...	2009
3	Empirical studies of a agile...	Tore Dyba * *, Torgeir Dings...	2009
4	Guidelines for performing S...	Barbara Kitchenham	2007
5	Motivation in Software Engi...	Sarah Beecham, Nathan Badd...	2007
6	Methods for the thematic s...	James Thomas and Angela Há...	2008
7	Systematic review of the us...	David Fone, Sandra Hollinghu...	2003

Figura 2. Tela do protótipo da ferramenta – funcionalidade que apresenta uma lista de estudos primários selecionados.

Com a definição dos estudos baseado no título, os próximos passos são a seleção dos estudos por resumo e por texto completo. Estas etapas não serão feitas automaticamente pela ferramenta, uma vez que é necessário que o usuário leia o conteúdo para fazer uma análise da relevância do estudo para a pesquisa. No entanto, apesar de ser uma análise manual a ferramenta SysReTool proverá auxílio durante todo o processo da revisão. Por exemplo, o gerenciamento dos estudos selecionados é feito pela ferramenta, onde o usuário poderá clicar no título do artigo e o resumo será exibido. Após a leitura do resumo, o usuário decidirá se aquele estudo é relevante. Caso não seja, ele será excluído. O mesmo processo será feito na última etapa, onde será exibido o texto completo e o usuário decidirá a relevância do estudo para a pesquisa.

Ao término, espera-se que tenha restado no conjunto de estudos selecionados apenas aqueles que o usuário julga serem relevantes à pesquisa em questão.

4. Cenário de execução da ferramenta

De modo a retratar o cenário de execução da ferramenta, a Figura 3 apresenta os passos que serão seguidos pela SysReTool para execução da revisão. Os dados apresentados nesta figura são fictícios, apenas para retratar a execução, em que é definida a busca e a seleção dos estudos, explicitando passo a passo a quantidade de artigos obtidos após cada passo que resulta na inclusão ou exclusão de artigos.

Como apresentado na Figura 3, existem estágios que são automatizados e outros manuais. As três primeiras etapas são executadas automaticamente pela ferramenta:

- **Etapa 1:** *Selecionar estudos por título:* a busca inicial retornou 1663 estudos. Após excluir trabalhos com títulos que claramente estavam fora do escopo restaram 99 estudos que continuaram a ser analisados.

- **Etapa 2:** *Eliminar estudos repetidos:* considerando que, havendo dois trabalhos semelhantes, o mais completo foi considerado, foram deletados 12 dos 99 estudos da etapa anterior, permanecendo 87 estudos para avaliação;
- **Etapa 3:** *Eliminar estudos por critérios definidos:* os 87 estudos foram submetidos a critérios definidos pelo usuário afim de obter uma pesquisa condizente com o perfil necessário. Restaram 48 estudos relevantes.

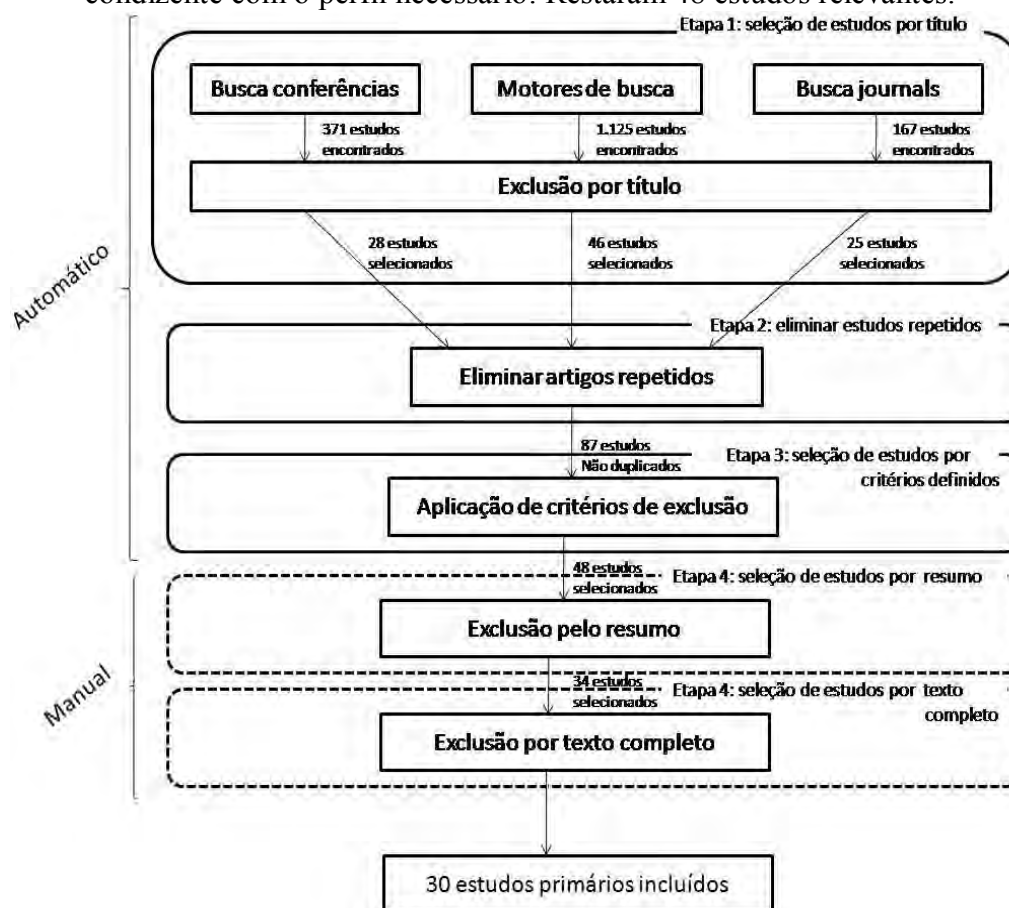


Figura 3. Passos definidos pela busca e seleção de trabalhos (adap. de Lobato (2012)).

A ferramenta armazenou uma lista com os 48 estudos relevantes obtidos após as etapas que executam o processo de eliminação e exclusão automaticamente. As próximas etapas são realizadas em modo manual, ainda com auxílio da ferramenta, que continua a armazenar a lista de estudos relevantes. As etapas manuais são:

- **Etapa 4:** *Selecionar estudos por resumo:* no exemplo, alguns estudos apresentavam informações superficiais sobre o assunto. Estes foram excluídos pelo usuário, restando 34 estudos.
- **Etapa 5:** *Selecionar estudos por texto completo:* o revisor considerou que 4, dos 34 estudos armazenados não eram relevantes e os excluiu. Logo, restaram 30 estudos ao final do processo de Revisão Sistemática.

5. Considerações Finais

Embora evidente os benefícios do uso de Revisões Sistemáticas para obtenção do estado-da-arte sobre uma determinada área de conhecimento, percebe-se que o seu processo de

elaboração, condução, e relato é trabalhoso e propenso a falhas. Desta forma, o uso de uma ferramenta computacional faz-se necessário, para que se obtenha maior qualidade dos resultados, ao passo que há redução de esforço. Uma proposta foi publicada recentemente que objetiva prover auxílio a este processo (Hernandes et al., 2012). Com base na análise dos requisitos desta ferramenta, decidimos cobrir um conjunto de aspectos que a complementasse, embora, a princípio, não haja relação entre ambas. A SysReTool inclui características de busca automáticas, que acreditamos ser elemento primordial de aumento de eficiência.

Foi publicado recentemente um estudo que verifica a viabilidade da identificação de estudos relevantes através da análise da seção de *abstract* dos artigos, e pela análise completa do texto (Torres, 2012). Baseado neste estudo nota-se que o uso da análise léxica e sintática e de expressões regulares podem auxiliar o processo de identificação de estudos primários. A classificação de relevância do estudo pode depender de alguns atributos como frequência das palavras-chave, tamanho da frase e posição das seções. De acordo com os estudos, vários métodos podem ser usados, mas ainda estão longe de chegar a um resultado ideal.

Este artigo apresenta os conceitos fundamentais de Revisões Sistemáticas de Literatura, e o workflow de uma proposta ferramenta que objetiva otimizar a realização de Revisões Sistemáticas, por meio da automatização de parte do processo., Apresentamos um exemplo prático, na tentativa de demonstrar sua viabilidade.

6. Referências

- Hernandes, E. C. M. ;Zamboni, A. B. ; Thommazo A. ; Fabbri, S. C. P. F.Using GQM and TAM to evaluate StArt a tool that supports Systematic Review. CLEI Electronic Journal, v. 15, p. 13-25, 2012.
- Kitchenham. B.; Charters, S. Guidelines for performing systematic literature reviews in software engineering (v. 2.3).Technical report, Keele University, 2007.
- Lima,M. S.; Soares, B. G. O.; Bacaltchuk, J. Psiquiatria baseada em evidências. Revista Brasileira de Psiquiatria, 2000.
- Lobato, L. L.An Approach for Risk Management in Software Product Lines. Recife: Universidade Federal de Pernambuco, 2012.
- Pai,M.; McCulloch, M.; Gorman, J. D.; Pai, N.; Enanoria, W.; Kennedy, G.; Tharyan, P.; Colford, J. M. Systematic reviews and meta-analyses: an illustrated, step-by-step guide. The National Medical Journalof India. 2004.
- Sampaio, R. F. andMancini, M. C. Estudos de Revisão Sistemática: um guia para síntese criteriosa da evidência científica. Revista brasileira de fisioterapia, São Carlos, v. 11, n.1, p. 83-89, jan/fev 2007.
- Torres, J.A.S.; Cruzes, D. S.; Salvador, L. N. Automatic Results Identification in Software Engineering Papers. Is it possible?. In: 2012 12th International Conference on Computational Science and Its Applications, Salvador-BA, p. 108-112.
- Walia, G. S. and Carver, J. C.A systematic Literature Review To Identify and Classify Software Requirement Errors. Information & Software Technology, v. 51, n. 7, pp. 1087-1109. 2009.

Construção de Bases de Dados para Auxiliar a Avaliação de Sistemas de Detecção de Intrusos em uma Rede IEEE 802.11 com Criptografia WEP, WPA e WPA2 Habilitada

Douglas Willer Ferrari Luz Vilela¹, Ailton Akira Shinoda², Ed'Wilson Tavares Ferreira³, Ruy de Oliveira³, Valtemir Emerêncio do Nascimento³, Nelcilen Araújo⁴

¹Coordenação de Tecnologia da Informação- Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) – Cuiabá - MT – Brazil

²Departamento de Engenharia Elétrica - Universidade Estadual Paulista Júlio de Mesquita Filho (UNESP) - Ilha Solteira - SP – Brazil

³Departamento de Informática - Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) – Cuiabá -MT – Brazil

⁴Instituto de Computação- Universidade Federal de Mato Grosso (UFMT) – Cuiabá - MT – Brazil

douglas.vilela@blv.ifmt.edu.br, shinoda@dee.feis.unesp.br, {ruy, ed, valtemir}@cba.ifmt.edu.br, nelcilen@ic.ufmt.br

Abstract. *The growth in the use of wireless technology IEEE 802.11 has been taking large proportions over recent years and the security mechanisms implemented in the same have proved to be ineffective in protection against denial of service attacks. Thus, the use of IDS as helping to secure networks, has increased. However, in order for the training algorithm intrusion detection to be effective it is necessary to use a database representative of traffic in the wireless environment. Though, one of major challenges in intrusion detection systems in wireless networks is limited number of methodologies for building database. This paper details the construction of databases, in which the first database is represented from a similar scenario of a typical for domestic use using WEP and / or WPA enabled. The second uses a similar scenario to that of a corporative environment with WPA2 encryption enabled.*

Resumo. *O crescimento da utilização da tecnologia sem fio IEEE 802.11 tem tomado grandes proporções nos últimos anos e os mecanismos de segurança implantado na mesma têm se mostrado pouco eficazes no combate a ataques de negação de serviço. Sendo assim, a utilização de IDS, como auxílio à proteção de redes, tem aumentado. Entretanto, para que o treinamento do algoritmo de detecção de intrusos seja eficaz é necessária a utilização de uma base de dados representativa do tráfego no ambiente sem fio. No entanto, um dos grandes desafios em sistemas detectores de intrusão em redes sem fio é número reduzido de metodologias para construção de base de dados. Esse trabalho detalha a construção de bases de dados, nas quais a primeira base de dados é representada a partir da implementação de um cenário semelhante à de um ambiente de uso doméstico utilizando criptografia WEP e/ou WPA habilitadas. A segunda utiliza um cenário semelhante ao de um ambiente corporativo com criptografia WPA2 habilitada.*

Introdução

Nos últimos anos houve um forte crescimento na utilização de redes móveis. O padrão IEEE 802.11 vem se destacando, como uma das tecnologias sem fio mais utilizadas na atualidade. Com o avanço na utilização das redes WLANs (*Wireless Local Area Networks*) começaram a surgir questionamentos sobre a eficácia dos mecanismos de segurança empregados nas mesmas. O primeiro protocolo de segurança lançado pela IEEE 802.11 foi o WEP (*Wired Equivalence Privacy*) em 1999. Este se mostrou muito vulnerável a ataques de negação de serviço e ataques de força bruta devido à chave de criptografia ser reduzida, reutilizada e estática. Em 2003 houve a primeira atualização na ementa de segurança do IEEE 802.11, sendo assim, foi implantado o protocolo WPA (*Wifi Protected Access*) apenas como solução temporária, enquanto o desenvolvimento da nova ementa era finalizado. Em 2004 foi lançado o IEEE 802.11i. O protocolo de segurança IEEE 802.11i trouxe melhorias relevantes, porém continuou não fornecendo proteção aos quadros de gerenciamento e controle, assim, mantendo as redes WLANs expostas a ataques de negação de serviço [Linhares e Gonçalves, 2008]. Em 2009, foi implementado padrão de segurança denominado IEEE 802.11w, este por sua vez adiciona proteção em alguns quadros de gerenciamento, mas ainda não protegia os quadros de controle [Ahmad e Tadakamadla, 2011].

As falhas existentes nas ementas de segurança do padrão IEEE 802.11 são evidentes. Isso mostra a necessidade em agregar outro mecanismo de segurança nas WLANs para minimizar as ameaças que exploram estas vulnerabilidades. A utilização de um IDS (*Intrusion Detection System*) é parte da solução de segurança. A detecção de intrusos é o processo de monitoramento de eventos que ocorre em um sistema computacional ou rede. Para que o IDS tenha um nível de eficiência alta é necessária uma capacidade de classificação correta, e isto se dá através de bases de dados coletadas para que o modelo de detecção possa aprender. Há uma grande dificuldade em encontrar este tipo de bases de conhecimento para redes IEEE 802.11, a maioria das bases existentes são voltadas para redes cabeadas LAN (*Local Area Network*) ou geradas em ambientes não realísticos [Shiravi et al, 2011].

O objetivo do artigo é apresentar a construção de bases de conhecimento para modelos de detecção de intrusos. As bases foram geradas a partir da implementação de uma rede IEEE 802.11 com criptografia WEP, WPA e WPA2 (*Wifi Protected Access 2*) habilitada em ambiente realístico. Em seguida, foram coletadas amostras de tráfego normal e tráfego anômalo da rede. Os ataques empregados para coleta do tráfego anômalo são do tipo negação de serviço, que tem como finalidade indisponibilizar os recursos e serviços da rede.

Este artigo está organizado em seções. A seção 2 apresenta os trabalhos relacionados encontrados na literatura. A seção 3 aborda como foram montados os experimentos para construção das bases de dados. A seção 4 apresenta as considerações finais e trabalhos futuros.

2. Trabalhos Relacionados

Em [Bicakci e Tavli, 2008], os autores fazem uma revisão teórica sobre ataques de negação de serviço que exploram vulnerabilidades da camada física e MAC das redes IEEE 802.11, com criptografia WEP, WPA e WPA2 habilitada. Ficaram evidentes as fragilidades apresentadas nos quadros de gerenciamento e controle, quando

submetidos a ataques de disponibilidade de serviço. Apesar de o trabalho explorar as vulnerabilidades não foi desenvolvido nenhum mecanismo para auxiliar na proteção das WLANs.

Na abordagem apresentada por [Shiravi et al, 2011], os autores relatam a escassez de bases de detecção de intrusos públicas. Apesar de existirem importantes contribuições como a da DARPA e KDD99, sua precisão e capacidades vêm sendo criticadas, devido ao fato de não terem sido construídas em ambientes reais. A fim de superar estas deficiências foi feita uma abordagem para gerar uma base de dados em um ambiente real, voltado para redes cabeadas, para analisar, testar e avaliar os sistemas de detecção de intrusos. Mesmo sendo construída uma base em ambiente real, ela não contém características específicas das redes sem fio, podendo influenciar na capacidade de detecção do IDS sem fio.

A proposta de [El-Khatib 2010], é gerar uma base de dados obtidos de uma rede sem fio com criptografia WEP e WPA habilitada. O autor menciona que para ter um melhor desempenho, precisão e confiabilidade no sistema de detecção de intrusos sem fio, é necessário extrair os cabeçalhos dos quadros para construção do algoritmo de detecção de intrusos. O trabalho foi realizado em um ambiente sem fio, mas não contempla ataques no padrão IEEE 802.11i.

A construção das bases de dados propostas neste artigo será gerada a partir de uma topologia de rede IEEE 802.11, aplicada em ambiente realístico, com criptografia WEP, WPA e WPA2 habilitada.

3. Metodologia

Os experimentos realizados na rede WLAN, para captura de pacotes e construção das bases de dados, são descritos nessa seção. Para geração da base de dados foram construídos dois cenários reais em ambientes controlados. A subseção 3.1 detalha o primeiro cenário que apresenta características semelhantes à de uma rede utilizada em um ambiente doméstico. No entanto, a subseção 3.2 apresenta o segundo cenário com características mais robustas semelhantes a uma rede utilizada em ambientes corporativos.

3.1. Construção da Base de Dados para Detecção de Intrusos em uma Rede IEEE 802.11 com Criptografia WEP/WPA Habilitada

Para construção da primeira base de dados, como apresenta a Figura 1, foi montada uma topologia de rede sem fio em um ambiente controlado, com características semelhantes a uma rede de uso doméstico, os seguintes mecanismos de segurança foram habilitados: criptografia WEP e WPA.

A topologia de rede foi constituída por: um ponto de acesso (AP), três estações sem fio. A estação tipo 1 inoculava tráfego de rede normal do tipo (HTTP, FTP). A estação tipo 2 empregou o uso da ferramenta Airplay [Aircrack, 2011] para executar os seguintes ataques de negação de serviço de maneira simultânea (*Chopchop*, *deautenticação*, *fragmentação* e *duração*). A estação do tipo 3 empregou o uso da ferramenta Wireshark [Wireshark, 2011] para executar a coleta das amostras e a ferramenta Tshark [Tshark, 2011] para pré-processar os dados.



Figura 1 – Topologia de rede WLAN aplicada no cenário 1.

Fonte: Elaborado pelo próprio autor.

Foram coletadas amostras de dados para construção da base conhecimento que será utilizada em treinamento de algoritmos de detecção de intrusos. É necessário executar um pré-processamento dos dados, para extrair apenas os campos do cabeçalho MAC dos quadros de gerenciamento, os respectivos campos são: *protocol version*, *type*, *subtype*, *to DS*, *from DS*, *more fragment*, *retry*, *power management*, *more data*, *WEP*, *order*, *duration*, *address1*, *address2*, *address3* e *sequence control*. O pré-processamento se faz necessário, pois os ataques empregados nesse experimento exploram vulnerabilidades dos quadros de gerenciamento. Portanto a base de conhecimento a ser aplicada na avaliação do classificador deve ser composta pelos quadros de gerenciamento da rede WLAN utilizada nos experimentos, em que a rede esteve sob condição “normal”, sem ataques, e sob os ataques *chopchop*, deautenticação, duração e fragmentação.

Esses ataques são do tipo DoS (*Denial of Service*) e exploram vulnerabilidades nos quadros de gerenciamento que afetam a disponibilidade dos recursos e serviços das redes IEEE 802.11 com criptografia pré-RSN. Os ataques de duração e deautenticação prejudicam a capacidade da estação base de gerenciar o acesso à infraestrutura da rede [Bellardo e Savage, 2003]. Os ataques de *chopchop* e fragmentação exploram as fragilidades dos mecanismos de criptografia (WEP e WPA) para deixar o serviço da rede indisponível [Bittau et al, 2006].

A base de dados foi criada utilizando captura por espaço amostral ao invés de captura por dias devido à restrição de recursos e mão de obra. As amostras foram divididas em três conjuntos: treinamento, validação e teste. A Tabela 1 apresenta a distribuição das amostras nos três conjuntos da base de dados e nas categorias de reconhecimento analisadas nos estudos.

A base é formada por 24200 amostras, estes valores foram reproduzidos a partir da base de conhecimento gerada no trabalho desenvolvido por [El-Khatib, 2010] que utiliza espaço amostral para captura dos dados, é constituído da seguinte forma: 9600 amostras conjunto de treinamento, sendo 6000 de tráfego normal e 3600 amostras de tráfego anômalo. O conjunto de validação é formado por 6400 amostras, sendo 4000 amostras de tráfego normal e 2400 amostras de tráfego anômalo. O conjunto de teste é

formado por 8200 amostras, sendo 5000 amostras de tráfego normal e 3200 amostras de tráfego anômalo.

Conjunto das bases de dados

	Treinamento	Validação	Teste
Normal	6000	4000	5000
ChopChop	900	600	800
Deautenticação	900	600	800
Duração	900	600	800
Fragmentação	900	600	800
Total de Amostras	9600	6400	8200

Tabela 1 – Distribuição das amostras nos conjuntos da base de dados cenário 1.

Fonte: Adaptado de [El-Khatib, 2010].

3.2. Construção da Base de Dados para Detecção de Intrusos em uma Rede IEEE 802.11 com Criptografia WPA2 Habilitada

Para construção da segunda base de dados para detecção de intrusos foi montada uma rede IEEE 802.11, com criptografia WPA2 habilitada, associação segura e mecanismo de autenticação 802.1x (Park et al, 2012). Este tipo de cenário é semelhante ao utilizado em ambientes corporativos, cuja topologia é representada na Figura 2.

A topologia de rede foi constituída por: cinco estações sem fio, dois pontos de acesso (AP) e um servidor RADIUS. As estações tipo 1,2 e 3 injetavam tráfego normal na rede (HTTP, FTP). A estação tipo 4 utilizou a ferramenta Airplay [Aircrack, 2011], Hping3 e Fake AP [Fakeap, 2007] para realizar os seguintes ataques aleatoriamente, (autenticação falsa, *synflooding*, de autenticação e AP falso). A estação tipo 5 utilizou a ferramenta Wireshark [Wireshark, 2011] para realizar a coleta e a ferramenta Tshark [Tshark, 2011] para realizar o pré-processar os dados. Servidor RADIUS realizava o controle de acesso a rede.

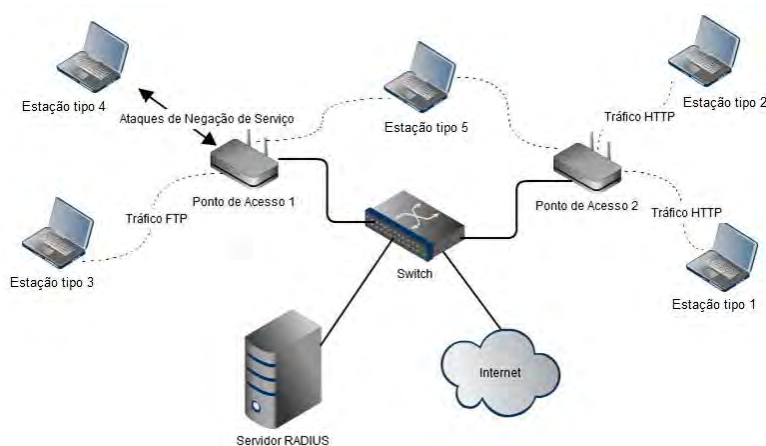


Figura 2 – Topologia de rede WLAN aplicada na geração da base de dados WPA2 habilitada.

Fonte: (Araújo, 2012).

Foi realizado um pré-processamento nos dados capturados para se extrair somente os campos do cabeçalho MAC (*protocol version, type, subtype, to DS, from DS, more fragment, retry, Power managment, more data, WEP, order, duration, address1, address2, address3 e sequence control*). A rede esteve sob condição “normal”, sem ataques, e sob os ataques autenticação falsa, *synflooding*, deautenticação e AP falso.

A base de dados gerada pelos dados coletados nas experiências segue o método de particionamento de dados *holdout*, desenvolvido por [Smith, 1994], que aloca o espaço amostral dos registros capturados numa proporção de 75% e 25%, respectivamente, nas bases de treinamento e teste.

Conjunto das bases de dados		
	Treinamento	Teste
Normal	4500	1500
Deautenticação	750	250
Autenticação Falsa	750	250
AP Falso	750	250
SynFlooding	750	250
Total de Amostras	7500	2500

Tabela 2 - Distribuição das amostras coletadas nos conjuntos da base de dados cenário 2.

Fonte: Adaptado de [Bishop, 1995].

4. Considerações finais e Trabalhos Futuros

Neste trabalho foram geradas duas bases de conhecimento que possuem potencial no treinamento de algoritmos de detecção de intrusos em redes IEEE 802.11. Isso gera a possibilidade do desenvolvimento de novos modelos de detecção utilizando as características das duas bases propostas. A intenção é que seja explorado o potencial das bases.

A contribuição do trabalho é a construção das bases de dados, que serão utilizadas para ajudar no treinamento de IDS para redes sem fio padrão IEEE 802.11, focando principalmente no problema dos ataques de negação de serviço. Além disso, espera-se que o trabalho estimule uma série de novos estudos na construção de bases de conhecimento por parte de pesquisadores da área.

Conclui-se que muito trabalho ainda pode ser feito. Como a implementação de novos de ataques de negação de serviço. A construção de uma nova topologia de rede, utilizando mecanismo de segurança 802.11w.

Referências

Aircrack (2011), <http://www.aircrack-ng.org/>.

- Ahmad, M. S. e Tadakamadla, S. (2011): “security evaluation of IEEE 802.11w specification”, Proceedings of the fourth ACM conference on Wireless network security. p. 53 - 58.
- Araújo, N., Oliveira, R., Shinoda, A., A., Ferreira, E., T. e Nascimento, V. E. (2012), “A Avaliação do Classificador ARTMAP Fuzzy em Redes 802.11 com Criptografia Pré-Robust Security Network (WEP e WPA)”, Anais do XII Simpósio Brasileiro em Segurança da Informação e de Sistemas Computacionais, p. 310-316.
- Bellardo, J. e Savage, S. (2003) “802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions”, Proceedings of the 12th Conference on USENIX Security Symposium, vol. 12, p. 15-28.
- Bicakci, K. e Tavli, B. (2009) “Denial-of-Service attacks and countermeasures in IEEE 802.11 wireless networks”. Computer Standards & Interfaces, vol. 31, p. 931-941.
- Bishop, C. M. (1995), Neural Networks for Pattern Recognition. Oxford University Press, 1st edition.
- Bittau, A., Handley, M. e Lackey, J. (2006) “The Final Nail in WEP's Coffin”, Proceedings of the 2006 IEEE Symposium on Security and Privacy, p. 386-400.
- El-Khatib, K. (2010) “Impact of Feature Reduction on the Efficiency of Wireless Intrusion Detection Systems”, IEEE Transactions on Parallel and Distributed Systems, vol. 21, n. 8, p. 1143-1149.
- Guenoun, M., Lbekkouri, A. e El-Khatib, K. (2008). “Optimizing the feature set of wireless intrusion detection systems”. International Journal Of Computer Science And Network Security, S.I., p. 127-131.
- IEEE Std. 802.11 (1999). “Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications”, p. 1-512.
- IEEE Std. 802.11i (2004). “Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY)”, p. 175.
- IEEE Std. 802.1X (2001). “IEEE Standard for local and metropolitan area networks - port-based network access control”.
- Linhares, A. G. e Gonçalves, P. A. D. S (2008). “Uma Análise dos Mecanismos de Segurança de Redes IEEE 802.11: WEP, WPA, WPA2 e IEEE 802.11w”. Universidade Federal de Pernambuco, p. 17.
- Park, K., Y., Kim, Y., S. e Kim, Y. (2012). “Security Enhanced IEEE 802.1x Authentication Method for WLAN Mobile Router”, Advanced Communication Technology (ICACT), 14th International Conference on. p. 549 – 553.
- Shiravi, A., Shiravi, H., Tavallae, M. e Ghorbani, A., A. (2012), “Toward Developing a Systematic Approach to Generate Benchmark Datasets for Intrusion Detection”, Computers & Security, vol. 31, p. 357 – 374.
- Smith, P., A. (1994). Autocorrelation in logistic regression modeling of species. Global Ecology and Biogeography letters, 4, p. 47-61
- Welch, D. e Lanthrop, S. (2003). “Wireless security threat taxonomy”, Proceedings of the 2003 IEEE workshop on information assurance. p. 76 - 83.
- Wireshark (2011), <http://www.wireshark.org/>.

Aplicativo classificador de imagens de notas de real para plataforma android

Lais C. R. da Silva Lopes¹, Mikaely C. Silva¹,
Allan K. Lopes¹, Fabrizzio A. A. de Melo N. Soares²

¹Instituto Federal Goiano – Campus Urutaí – Departamento de Informática
Rodovia Geraldo Nascimento, Km 2,5 – Urutaí – GO – Brasil

²Universidade Federal de Goiás – Instituto de Informática
Campus Samambaia, Bloco IMF 1 – Goiânia – GO – Brasil

{allankardeci, laiscandidolopes, mikaelycruciol42}@gmail.com, fabrizzio@inf.ufg.br

Abstract. *This paper proposes a system to classify images of Real notes using the data decorrelation technique known as Principal Component Analysis (PCA) and an Artificial Neural Network (ANN). The PCA was used during the preprocessing of images to reduce the size of the data. The new representation of the images, created by the PCA, served as input to an ANN, the type Multilayer Perceptron (MLP) make the classification. After the trained classifier, the state of the objects that compose are preserved and ported to an application developed for the Android platform. The goal is to assist blind people to identify the brazilian banknotes. The results demonstrated that the system was efficient, reaching a hit percentage of almost 100% for all of the images used.*

Resumo. *Este trabalho propõe um sistema classificador de imagens de notas de Real com o uso da técnica de decorrelação de dados, conhecida como Análise das Componentes Principais (PCA), e uma Rede Neural Artificial (RNA). O PCA foi utilizado durante a fase de pré-processamento das imagens para diminuir a dimensão dos dados. A nova representação das imagens, criada pelo PCA, serviu de entrada para uma RNA, do tipo Multilayer Perceptron (MLP), fazer a classificação. Depois de treinado o classificador, o estado dos objetos que o compõe são preservados e portados para um aplicativo desenvolvido para a plataforma Android. O objetivo é auxiliar pessoas cegas a identificar as cédulas monetárias brasileiras. Os resultados obtidos demonstram que o sistema foi eficiente, atingindo uma porcentagem de acerto de quase 100% para o conjunto de imagens utilizado.*

1. Introdução

O Brasil possui uma população significativa de deficientes visuais. De acordo com o censo realizado em 2010 pelo IBGE¹, cerca de 18,8% da população do Brasil possui algum tipo de dificuldade visual [IBGE 2012]. Embora essa população seja expressiva ainda há uma carência muito grande de adaptações para deficientes visuais que enfrentam uma série de dificuldades quando interagem com o ambiente à sua volta. A quantidade de informação codificada visualmente torna essa interação ainda mais desafiadora. Uma

¹Instituto Brasileiro de Geografia e Estatística

dificuldade muito comum de uma pessoa cega é reconhecer o valor das notas de dinheiro que está segurando.

No Brasil, a moeda corrente é o Real composta pelas notas de R\$ 1, 2, 5, 10, 20, 50 e 100. Ressalta-se que, desde a introdução do Real as notas sofreram diversas substituições. Essas substituições foram por causa da evolução de mecanismos de segurança, introdução de elementos históricos ou culturais e em alguns casos acréscimo de elementos para apoio aos deficientes visuais. A nota de R\$ 1, que atualmente está fora de circulação, de 1994 a 2003 teve 20 versões [Banco Central do Brasil 2013]. Cada substituição contribui para a confusão dos deficientes visuais, que podem levar bastante tempo para se familiarizarem com as novas notas.

Diversos trabalhos vem sendo pesquisados para reconhecimento de notas. Em [Hasanuzzaman et al. 2011] foi pesquisado o reconhecimento de notas de dólar utilizando um modelo baseado em componentes a partir da tecnologia *Speeded Up Robust Features* (SURF). No trabalho de [Toytman and Thambidurai 2011] também foi utilizado o reconhecimento de notas de dólar com descritores SURF, porém, com o objetivo de utilização em dispositivos móveis baseados na plataforma Android. [Hinwood et al. 2006] desenvolveu um equipamento para reconhecimento de notas de dinheiro que internamente possui os dispositivos BlindNoteReader, NoteTeller 2 e outros. No trabalho de [Sevcli et al. 2002] foi realizado o reconhecimento de notas utilizando o algoritmo de aprendizado indutivo RULES-3. Em [Aoba et al. 2003] foram experimentadas as RNA's para reconhecimento de notas de Euro com características extraídas de imagens e infravermelho. [Chae et al. 2009] utilizou informações digitalizadas de notas de dinheiro expostas a ultravioleta e comparou com mecanismos contra falsificação previstos nas notas.

Neste trabalho é proposto um sistema para reconhecimento de notas de Real na plataforma Android [Google 2013] usando PCA e uma RNA do tipo MLP. O objetivo é auxiliar pessoas com algum tipo de dificuldade para enxergar a identificarem as notas da moeda brasileira. Ressalta-se que o sistema proposto neste trabalho possui apenas a habilidade de identificar o valor das notas e não constatar sua legitimidade.

2. Revisão sobre análise das componentes principais

O método de descortelamento dos dados, Principal Component Analysis (PCA) [Hotelling 1933] ou Karhunen-Loève Transform (KLT) [Karhunen 1947] e [Loève 1977] tem como objetivo a obtenção de um pequeno número de combinações lineares, componentes principais, de um conjunto de variáveis que retenham o máximo possível da informação contida nas variáveis originais. As componentes são extraídas na ordem da mais explicativa para a menos explicativa, sendo possível descartar as menos importantes, compactando dessa forma a informação original.

Aplicações com imagens que utilizam a técnica geralmente constroem uma matriz formada por um conjunto de vetores $X = \{X_n R^d \mid n = 1, \dots, N\}$ em que as linhas de X são formadas pelos dados de imagens concatenadas, com d sendo o produto entre a altura e largura de uma imagem. Em seguida, calcula-se o vetor média de X conforme equação 1.

$$E(X) = \frac{1}{N} \sum_{n=1}^N X_n \quad (1)$$

O vetor média do conjunto representa tudo aquilo que é comum a todas as imagens. Após subtrair cada elemento de X da média obtém-se uma nova matriz que contém somente a variação de cada imagem em torno da média, para isto cada linha de X deve ser subtraída de $E(X)$, conforme equação 2.

$$\bar{X}_n = \{X_n - E(X), n = 1, \dots, N\} \quad (2)$$

Calcula-se a matriz de covariância $M = \bar{X} \times \bar{X}^T$, encontram-se os autovetores Q e autovalores V da matriz M , para então, recalculamos a matriz de autovetores que seja ortogonal e ortonormal, ou seja, as componentes principais, conforme equação 3.

$$Q^* = \frac{\bar{X}^T \times Q \times (V)^{-0.5}}{\sqrt{N-1}} \quad (3)$$

Finalmente, projeta-se a matriz $\bar{X}$ na matriz transformação Q^* para se obter uma nova representação das imagens. Pode-se ordenar os autovetores da matriz Q^* de acordo com a significância dada pela matriz de autovalores V e assim eliminar os autovetores de menor representatividade.

Para aplicar o PCA em uma imagem de teste, afim de classificá-la, é necessário projetá-la na matriz transformação Q^* . Seja x_{teste} a matriz que representa a imagem de teste; então x_{teste}^T será essa matriz concatenada e $\bar{x}_{teste}^T$ será essa matriz concatenada subtraída da média global $E(x)$, conforme equação 4.

$$\bar{x}_{teste}^T = x_{teste}^T - E(x) \quad (4)$$

Projeta-se $\bar{x}_{teste}^T$ na matriz transformação, conforme equação 5, e utiliza-se uma regra de decisão afim de se classificar a imagem de teste.

$$\bar{x}_{pca} = \bar{x}_{teste}^T \times Q^* \quad (5)$$

3. Modelo proposto

Serão detalhadas a seguir as principais etapas de construção do classificador de notas de Real usando PCA e RNA.

3.1. Digitalização das notas para treinamento e teste

Para o treinamento e teste do sistema foram digitalizadas diversas notas de Real. Foram digitalizadas notas de 2, 5, 10, 20, 50 e 100 reais. As notas de 100 reais foram utilizadas em suas duas versões existentes. As notas foram digitalizadas em frente e verso sob um fundo branco. De acordo com as condições de conservação e semelhança de cada nota, algumas notas não tiveram ambas as faces digitalizadas.

Para a captura das imagens das notas foram utilizadas as câmeras RGB primárias dos celulares Samsung modelos Galaxy Y GT S6102B e Galaxy Pocket S5300, com 3 e 2 megapixels de resolução, respectivamente. As digitalizações foram realizadas manualmente e em diferentes condições de luminosidade, distância e posição. A figura 1 mostra exemplos das imagens de notas de Real capturadas e a tabela 1 apresenta a quantidade de imagens digitalizadas de cada nota.

Tabela 1. Quantitativo de amostras digitalizadas de cada nota

Nota	Qtd. Frente	Qtd. Verso	Total
2	38	38	76
5	35	27	62
10	24	34	58
20	31	32	63
50	35	42	77
100 v1	41	47	88
100 v2	35	41	76
Total	239	261	500



Figura 1. Exemplos de imagens de frente e verso das notas de 2, 10 e 50

3.2. Pré-processamento

Esta fase pode ser dividida em dois passos, no primeiro as imagens obtidas através da atividade descrita na subseção 3.1, foram convertidas em imagens de 256 tons de cinza e redimensionadas para a resolução de 300×140 pixels de largura e altura, respectivamente. Em seguida, dividiu-se as imagens em 2 conjuntos, um para treinamento composto por 90% das imagens e outro para teste com os 10% restantes. Esses conjuntos contêm representantes de todas as classes, por exemplo, para a nota de 2 foram capturadas 38 imagens da parte da frente e 38 do verso, neste caso, separou-se 90% das imagens da parte da frente e do verso para treinamento e 10% das imagens da parte da frente e do verso para teste.

O segundo passo refere-se à aplicação do PCA, descrito na seção 2, ao conjunto de treinamento, o resultado da aplicação do PCA foi utilizado como entrada para a RNA.

3.3. Arquitetura da rede neural artificial

A classificação das imagens foi feita por uma RNA, do tipo *Multilayer Perceptron*, treinada com o algoritmo *backpropagation* [Rumelhart et al. 1986], com *momentum* e taxa de aprendizagem adaptativa [Vogl et al. 1988], para a atualização dos pesos e bias dos neurônios da rede.

A RNA foi projetada com três camadas: uma de entrada, uma oculta e uma de saída. O número de neurônios na camada de entrada é igual ao tamanho dos vetores de entrada que representam as imagens. Essa camada não realiza nenhum processamento,

tem a função de apenas distribuir os sinais para a próxima camada de neurônios. Na camada oculta o número de neurônios, no experimento realizado, variou entre 3 e 60 e para a camada de saída foram necessários 6 neurônios, um por nota.

Os pesos e bias das conexões dos neurônios da rede foram inicializados de acordo com a regra de D. Nguyen e B. Widrow[Nguyen and Widrow 1990]. A figura 2 ilustra a RNA projetada para classificar as imagens. Em ambas as camadas, oculta e de saída, os neurônios utilizam a função de transferência tangente hiperbólica sigmóide.

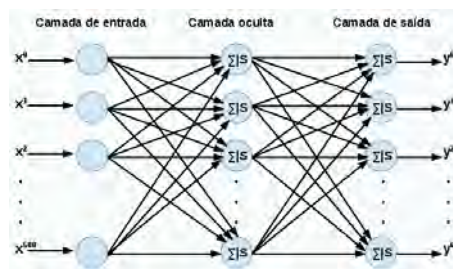


Figura 2. Representação da rede neural projetada

O classificador proposto foi codificado na linguagem de programação Java utilizando o paradigma da orientação a objetos. A figura 3 mostra as principais classes desenvolvidas. Afim de não poluir visualmente o diagrama somente os principais atributos e métodos são mostrados.



Figura 3. Principais classes do sistema proposto

3.4. Treinamento e teste da rna

Como explicado na subseção 3.2, as imagens capturadas foram divididas em dois conjuntos, o primeiro, utilizado durante o treinamento da RNA, composto por 90% das imagens e o segundo, utilizado durante a fase de teste, composto pelos os 10% restantes.

O algoritmo de treinamento da RNA foi executado 20 vezes com diferentes números de neurônios na camada oculta, adotou-se a estratégia de aumentar de 3 em 3 a quantidade de neurônios em cada execução, a primeira execução inicia-se com 3 neurônios. A tabela 2 mostra os parâmetros utilizados nas execuções do algoritmo de treinamento.

Antes do início do aprendizado da RNA aplicou-se o PCA ao conjunto de imagens de treino, conforme o que foi mencionado na seção 2 e na subseção 3.2. Após o treino

empregou-se as equações 4 e 5 a cada uma das imagens do conjunto de teste afim de verificar o desempenho da RNA.

Dois critérios de parada foram adotados nas sessões de treinamento da RNA. O treinamento termina caso o erro mínimo aceitável ou o número máximo de ciclos sejam atingidos. A fim de identificar a causa do término do treino foi adicionado um código para informar a causa de parada (1 - parada por atingir o número máximo de ciclos e 2 - parada por ter atingido o erro mínimo aceitável), conforme pode ser visto na tabela 3.

Tabela 2. Valores dos parâmetros utilizados nas execuções do algoritmo de treinamento da RNA

Parâmetro	Valor	Descrição
lr	0,01	Taxa de aprendizado inicial
inc	1,05	Proporção de incremento da taxa de aprendizado
dec	0,7	Proporção de decremento da taxa de aprendizado
momentum	0,9	Constante <i>momentum</i>
max_inc_erro	1,04	Valor máximo aceitável de aumento do erro médio quadrático
gradiente	0,0000001	Valor do erro global aceitável da rede
epocas	3000	Número máximo de ciclos de treinamento da RNA

3.5. Detalhes da Implementação do Aplicativo para Android

Após o treinamento e teste da RNA, em um computador *desktop*, momento em que teoricamente o classificador estaria pronto para uso, os objetos das classes, mostradas na figura 3, são salvos em disco e reutilizados posteriormente em uma aplicação desenvolvida para Android sem a necessidade de execução de todos os outros passos.

A aplicação foi desenvolvida utilizando o ambiente de desenvolvimento integrado Eclipse versão 4.2.1 codinome Juno, juntamente com o plugin *Android Development Tools* (ADT). Os objetos salvos em disco foram copiados para a memória permanente de um aparelho celular modelo Galaxy Samsung SII GT-I9100. A aplicação desenvolvida realiza as seguintes operações: 1^o - captura a imagem de teste através da câmera do aparelho, 2^o - converte a imagem em 256 tons de cinza e a redimensiona para a resolução de 300×140 de largura e altura, 3^o - aplica o PCA à imagem obtida, conforme equações 4 e 5, 4^o - carrega para a memória principal os objetos armazenados em sua memória permanente e 5^o - classifica a imagem utilizando a RNA.

Para facilitar a utilização do aplicativo foi disponibilizado um atalho na área de trabalho do dispositivo. Uma vez iniciado o programa, o usuário deve apenas direcionar a câmera para a nota, clicar no único botão disponível na interface e esperar pela resposta.

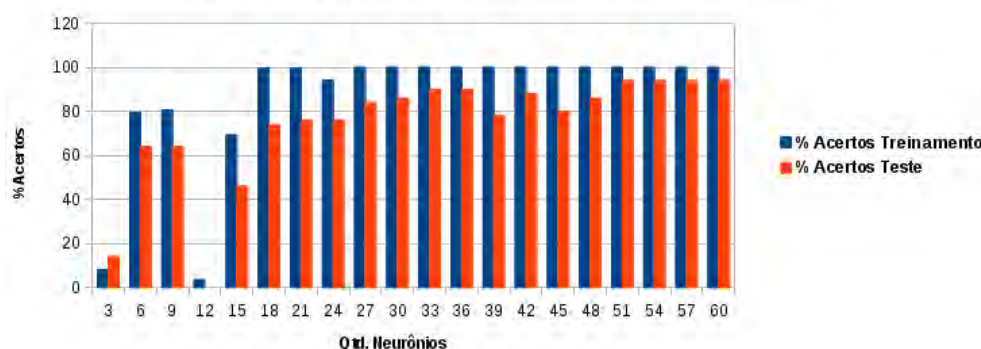
4. Resultados

Na tabela 3 observa-se que a partir de 18 neurônios na camada oculta a RNA já é capaz de fazer uma boa classificação das imagens e que a maioria das execuções do algoritmo de treinamento terminou antes da quantidade máxima de ciclos estipulada (3000), ou seja, convergindo antes para o erro mínimo aceitável de 0,0000001.

O gráfico da figura 4 mostra que a partir de uma certa quantidade de neurônios na camada oculta o desempenho da RNA torna-se estável, não influenciando muito aumentar o número de neurônios nessa camada. Os melhores resultados foram obtidos nas execuções 17, 18, 19 e 20. Nestes casos a RNA errou apenas 3 das 50 amostras do conjunto de teste. Acredita-se que estes resultados podem ser incrementados, ajustando os parâmetros do algoritmo para que convirja mais lentamente. Na maioria das execuções o algoritmo convergiu antes dos 1500 ciclos.

Tabela 3. Porcentagem de acertos da RNA para os conjuntos de treinamento e teste em cada execução

Exec.	Cód. de Parada	Qtd. Neurônios	% Acertos Treinamento	% Acertos Teste
1	1	3	8	14
2	1	6	79,5	64
3	1	9	80,6	64
4	2	12	3,5	0
5	1	15	69,3	46
6	1	18	99,7	74
7	1	21	99,7	76
8	1	24	94,2	76
9	2	27	100	84
10	2	30	100	86
11	2	33	100	90
12	2	36	100	90
13	2	39	100	78
14	2	42	100	88
15	2	45	100	80
16	2	48	100	86
17	2	51	100	94
18	2	54	100	94
19	2	57	100	94
20	2	60	100	94

Relação entre a quantidade de neurônios e o desempenho da RNA**Figura 4. Quantidade de neurônios e o desempenho da RNA**

5. Conclusão

O PCA foi eficiente na redução e caracterização das imagens de notas Real compactando significativamente o tamanho das imagens. Apesar de perda de informação envolvida no processo de obtenção das componentes principais a RNA conseguiu fazer um bom trabalho, aprendendo e generalizando bem os padrões apresentados. Talvez seja necessário modificações no algoritmo de treinamento para que o seu desempenho seja ainda melhor, já que 1500 ciclos de treinamento, em média, talvez não seja o suficiente para a RNA classificar corretamente todas as imagens.

Apesar dos resultados serem positivos é necessário cautela pois as imagens, de certa forma, foram capturadas em um ambiente controlado, todas estão na frente de um fundo branco. Pretende-se capturar outras imagens em ambientes mais reais para apresentar para o classificador já treinado afim de observar como ele se comporta, caso se saia mal, pode-se treinar a RNA novamente.

Em trabalhos futuros pretende-se adicionar novas funcionalidades ao aplicativo desenvolvido afim de torná-lo ainda mais fácil de se utilizar além da realização de estudos comparativos entre o classificador apresentado e outras técnicas de classificação de padrões com o objetivo de desenvolver algo ainda mais eficiente.

References

- Aoba, M., Kikushi, T., and Takefuji, Y. (2003). Euro Banknote Recognition System Using a Three-layered perceptron and RBF Networks. *IPJS Transactions on Mathematical Modeling and Its Application*, 44(SIG 7 (TOM 8)).
- Banco Central do Brasil (2013). Museu de Valores do Banco Central do Brasil.
- Chae, S.-H., Kim, J., and Pan, S. (2009). A Study on the Korean Banknote Recognition Using RGB and UV Information. In Ślęzak, D., Kim, T.-h., Chang, A.-C., Vasilakos, T., Li, M., and Sakurai, K., editors, *Communication and Networking*, volume 56 of *Communications in Computer and Information Science*, pages 477–484. Springer Berlin Heidelberg.
- Google (2013). Sítio oficial do sistema operacional android. www.android.com.
- Hasanuzzaman, F. M., Xiaodong, Y., and YingLi, T. (2011). Robust and effective component-based banknote recognition by SURF features. In *Wireless and Optical Communications Conference (WOCC), 2011 20th Annual*, pages 1–6.
- Hinwood, A., Preston, P., Suaning, G. J., and Lovell, N. H. (2006). Bank note recognition for the vision impaired. *Australasian Physical & Engineering Sciences in Medicine*, 29(2).
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *J. Educ. Psychol. Columbia University*, 24:417–441 e 498–520.
- IBGE (2012). Censo 2010. <http://www.censo2010.ibge.gov.br/resultados>.
- Karhunen, K. (1947). *Über lineare Methoden in der Wahrscheinlichkeitsrechnung*. Suomalaisen Tiedekatemian toimituksia: Mathematica-physica. Universitat Helsinki.
- Loève, M. (1977). *Probability Theory I*. Number 9780387902104 in Graduate Texts in Mathematics. Springer.
- Nguyen, D. and Widrow, B. (1990). Improving the learning speed of 2-layer neural networks by choosing initial values of the adaptive weights. In *Neural Networks, 1990., 1990 IJCNN International Joint Conference on*, volume 3, pages 21–26. IEEE.
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning internal representations by error propagation. *Mit Press*.
- Sevcli, M., Türkyilmaz, A., and Aksoy, M. S. (2002). Banknote recognition using inductive learning. Proceedings of Int. Conf. On Fuzzy Syst. And Soft Computational Intelligence in Management and Industrial Eng - 15. FSSCIMIE'02, pages 122–128.
- Toytman, I. and Thambidurai, J. (2011). Banknote recognition on Android platform.
- Vogl, T. P., Mangis, J. K., Rigler, J. K., Zink, W. T., and Alkon, D. L. (1988). Accelerating the convergence of the backpropagation method. *Biological Cybernetics*, 59:257–263.

Aplicação para dispositivos móveis utilizando tecnologias interativas: a Realidade Virtual e Aumentada aplicada ao estudo da anatomia humana

Lilian R. Mendes Paiva¹, Flávia G. Fernandes², Austeir José Barbosa², Camila Caroline de O. Moura², Luciene C. de Oliveira², Mylene L. Rodrigues²

¹Campus Avançado - Instituto Federal do Triângulo Mineiro (IFTM)
CEP: 38.408-343 – Uberlândia – MG – Brasil

²Instituto de Engenharia e Tecnologia – Universidade de Uberaba (UNIUBE)
CEP: 38.408-343 – Uberlândia – MG – Brasil

lilianpaiva@iftm.edu.br,
{flavia.fernandes92,auteir,camilah.kat}@gmail.com,
{luciene.oliveira,mylene.rodrigues}@uniube.br

Abstract. *This paper presents the proposal and the results a project whose goal is to develop a system to support learning of human anatomy, linking the areas of health and education. The methodology will use the resources of the Virtual and Augmented Reality (RVA) that enables the visualization, interaction and simulation of three-dimensional images. It is expected to contribute in this way in motivating students, through technological innovations, creating an educational environment pleasant and harmonious addition to continuing education of professionals and the routine of doctors, which would help to better inform patients about the functioning the human organism.*

Resumo. *Este artigo apresenta os resultados parciais de um projeto de pesquisa, cujo objetivo geral é o desenvolvimento de um sistema para apoio à aprendizagem da anatomia humana, unindo as áreas da saúde e educação. Como metodologia, serão utilizados os recursos da Realidade Virtual e Aumentada (RVA,) que possibilitam a visualização, interação e simulação tridimensional de imagens. Espera-se, contribuir desta forma na motivação de alunos, através das inovações tecnológicas, gerando um ambiente educacional agradável e harmonioso além da formação continuada dos profissionais da área e na rotina dos consultórios médicos, o que ajudaria a informar melhor os pacientes sobre o funcionamento do organismo humano.*

1. Introdução

Recentemente, observa-se que o desenvolvimento de tecnologias da informação vem auxiliando inúmeras práticas na área da saúde, em atividades como diagnóstico, terapia, gerenciamento e educação.

Dentro de tais práticas, encontra-se o apoio ao ensino da anatomia humana, campo que estuda grandes estruturas e sistemas do corpo humano. Atualmente, vivencia-se que nas aulas práticas em laboratórios os discentes apresentam distúrbios, como pesadelos, insônia, ansiedade, repulsa visual, depressão e outros. A morte, o formol e o odor juntos, instigam situações de repelência. Isso sugere casos de possíveis barreiras para o real aprendizado [Espírito, 1981].

Paralelamente, a aquisição de cadáveres se torna cada vez mais difícil, considerando a tradição desde os primórdios da sociedade, quando o cadáver é considerado algo sagrado, voltado ao culto religioso. Pesquisadores também apontam

quanto à dificuldade de alcançar os propósitos de ensino e aprendizagem, utilizando cadáveres descoloridos, firmes e elásticos, obtidos pela preservação em formalina a 10%, algumas vezes tão irreconhecíveis que não apresentam a mínima semelhança com o tecido original [Edelweiss, 1993].

Consequentemente, embora seja fundamental ao ensino da área de saúde, observa-se que, inúmeras vezes, os estudantes percebem a devida importância do estudo em anatomia humana somente quando estão vivenciando situações práticas ao lado de um leito ou de uma mesa operatória [Chopard, 1997].

Diante deste contexto, o ensino da anatomia humana, precisa ser repensado para corresponder às expectativas deste novo e atual momento. A inserção da tecnologia no ensino da anatomia humana poderá propiciar alternativas para dinamizar as aulas, favorecendo uma aprendizagem coerente com a realidade ética e social e com o mercado de trabalho que os estudantes dos cursos da área de saúde encontrarão após sua graduação [Martins, 2009; Quartiero, 1999; Lévy, 1993].

Neste campo, a Realidade Virtual e Aumentada (RVA), destaca-se pelas novas possibilidades de criação, modelagem, visualização, interação e simulação tridimensional de imagens, proporcionando interfaces avançadas capazes de gerar a imersão do usuário em ambientes com os quais pode interagir e explorar [Cardoso, 2007].

Desse modo, a aplicação da RVA na área da saúde, tem sido avaliada de forma intensiva nos últimos anos, com merecido destaque, pois representa novos desafios e potencialidades, com a inserção de informações complementares e/ou relevantes ao cenário real. Isoladamente, a RVA não atinge os objetivos esperados da complexa tarefa de ensino e aprendizagem. Porém, pode ser utilizada como ferramenta de apoio didático ou mesmo, empregada em diversos momentos do ensino-aprendizagem nas salas de aula presencial ou virtual. Nesta linha de raciocínio, esta pesquisa tem como objetivo propor uma ferramenta de RVA na área da aprendizagem da anatomia do corpo humano, unindo as áreas da saúde e educação.

Logo, os resultados motivarão o aprendizado dos alunos de maneira interativa e lúdica, tornando o aprendizado mais atrativo e dinâmico. Trazendo este recurso para as escolas, os alunos comprometer-se-iam mais com os estudos pela motivação com as inovações tecnológicas, o ambiente educacional tornar-se-ia mais agradável e harmonioso, o que incentivaria a formação continuada dos profissionais da área.

2. Fundamentos da Realidade Virtual e Aumentada

Realidade Virtual (RV) e Realidade Aumentada (RA) são duas áreas relacionadas com as novas gerações de interface do usuário, facilitando e potencializando as aplicações computacionais. Existem muitas definições para RVA, envolvendo diversos aspectos, a saber:

- **Realidade Virtual** é uma interface avançada para aplicações computacionais, que permite ao usuário navegar e interagir, em tempo real, com um ambiente tridimensional, usando dispositivos multisensoriais [Kirner e Siscoutto, 2007].

- **Realidade Aumentada** é a ampliação da percepção sensorial por meio de recursos computacionais, gerando a mistura dos mundos reais e virtuais, conectando ambientes completamente reais e virtuais [Lamounier e Cardoso, 2004].

A RV associa-se a uma interface homem-máquina poderosa, possibilitando ao usuário interação, navegação e imersão num ambiente tridimensional sintético, gerado pelo computador através de canais multissensoriais, podendo ser categorizada em:

- **RV imersiva** objetiva isolar o usuário por completo do mundo real, utilizando equipamentos como capacete de realidade virtual, luva de dados, rastreadores e fones de ouvido a fim de responder somente aos estímulos gerados pelo sistema computacional [Harma, 2003].
- **RV não-imersiva**, o usuário tem acesso ao ambiente virtual, sem se isolar do mundo real, isto é, através de dispositivos convencionais de computador (monitor e mouse) [Harma, 2003].

Um aplicativo de RV pode proporcionar uma sessão sob três formas diferentes: [Costa, 2009].

Tabela 1. Sessão de Realidade Virtual

Passiva	Proporciona ao usuário uma exploração do ambiente automática e sem interferência. A rota e os pontos de observação são explícitos e controlados exclusivamente pelo software. O usuário não tem controle algum, exceto talvez, para sair da sessão.
Exploratória	Proporciona uma exploração do ambiente dirigida pelo usuário. O participante pode escolher a rota e os pontos de observação, mas não pode interagir de outra forma com entidades contidas na cena.
Interativa	Proporciona uma exploração do ambiente dirigida pelo usuário e, além disso, as entidades virtuais do ambiente respondem e reagem às ações do participante.

A RV é uma particularização da Realidade Misturada, que consiste na sobreposição de objetos virtuais tridimensionais gerados por computador com um ambiente físico, realizada em tempo real, através de um dispositivo tecnológico [Reimann, 2005].

A Realidade Misturada (RM) pode receber duas denominações: RA, quando o ambiente principal ou predominante é o mundo real e, Virtualidade Aumentada, quando o ambiente principal ou predominante é o mundo virtual [Neto, 2006].

No caso da RA, cabe ressaltar que, os objetos virtuais são trazidos para o espaço do usuário, onde ele tem segurança e sabe como interagir, sem necessidade de treinamento. Desta forma, a RA pode ser definida como uma tecnologia através da qual se incrementa ou aumenta a visão que um utilizador tem do mundo real com a adição de imagens virtuais, usando técnicas de visão por computador e de Computação Gráfica/Realidade Virtual, resultando na sobreposição de objetos virtuais com o mundo real [Drab, 2005]. Além de permitir essa sobreposição de objetos virtuais no mundo real, a RA também permite o manuseio desses objetos com as próprias mãos, possibilitando que o usuário tenha uma interação atrativa e motivadora com o ambiente [Siscoutto, 2008]. No entanto, para que esses objetos sobrepostos sejam visualizados e manipulados, é necessária a utilização de algum software e dispositivo tecnológico.

Dessa maneira, as principais vantagens da utilização de técnicas de RV/RA para fins educacionais são as seguintes [Cardoso, 2007]:

- Motivação de estudantes e usuários de forma geral, baseada na experiência de 1ª pessoa vivenciada pelos mesmos;

- Grande poderio de ilustrar características e processos, em relação a outros meios multimídia;
- Permite visualizações de detalhes de objetos;
- Suporte às visualizações de objetos que estão a grandes distâncias, como um planeta ou um satélite;
- Permite experimentos virtuais, na falta de recursos, ou para fins de educação virtual interativa;
- Permite ao aprendiz refazer experimentos de forma atemporal, fora do âmbito de uma aula clássica;
- Possibilita interação, exigindo que cada participante se torne ativo dentro de um processo de visualização;
- Encoraja a criatividade, catalisando a experimentação;
- Provê igual oportunidade de comunicação para estudantes de culturas diferentes, a partir de representações;
- Ensina habilidades computacionais e de domínio de periféricos.

3. Metodologia

O processo de desenvolvimento de sistemas contempla antecipadamente a modelagem para análise através de métodos formais e gráficos, que abrangem de forma integral a representação dos sistemas. Dentro das práticas recomendadas pela programação orientada a objetos, este projeto será desenvolvido através de diversas fases, descritas a seguir [Bobbch, 2005]:

1. Fase de Concepção: documentação contendo modelos que contenham os requisitos da aplicação a ser construída baseada em Casos de Uso, utilizando a linguagem de modelagem UML (*Unified Modeling Language*).
2. Fase de Elaboração: documentação técnica para a arquitetura do sistema contendo diagramas como Modelo de Dados, Diagramas de Atividades, Diagramas de Estado, Modelo de Deployment e Implantação.
3. Fase de Construção: desenvolvimento do software para RVA utilizando dispositivos móveis e interação natural por meio de reconhecimento de gestos.
4. Fase de Testes: plano de testes e realização de testes integrados na aplicação desenvolvida.

Em relação aos aspectos metodológicos e tecnológicos, será utilizada RVA por meio do desenvolvimento de ambientes virtuais, incluindo interações e animações, com uso das linguagens de programação Java e Javascript, a linguagem de marcação HTML (*HyperText Markup Language* - Linguagem de Marcação de Hipertexto) para estruturação dos conteúdos, a linguagem CSS (*Cascading Style Sheets* - Folha de Estilos em Cascata) para formatação dos conteúdos estruturados, além do ambiente de desenvolvimento Eclipse, conectado ao emulador do sistema Android, e também a ferramenta de modelagem gráfica Blender, a fim de contemplar cálculos e comandos, visando criar para o usuário a possibilidade de interagir com ambientes virtuais atrativos, que facilitem a aprendizagem significativa de conceitos teóricos e práticos. Também utilizará o sistema Android para dispositivos móveis na realização de testes e para sua utilização prática e efetiva [Bodoff, 2005].

O emulador Android, também conhecido como Android Virtual Device (AVD), é uma reconhecimento de características de smartphones usando a versão do Android™ escolhida (seja a 2.1, 2.2 ou qualquer outra). Com ele, são carregados a imagem do sistema e características, onde o usuário pode realizar testes no projeto [Kanno, 2011].

O Android SDK é um aplicativo instalado no software Eclipse que permite que os desenvolvedores elaborem as aplicações a partir de um dispositivo virtual para os aparelhos de celular e tablet, desde jogos a utilitários que façam uso das funções oferecidas pelos aparelhos, como touchscreen, telefonia GSM, Câmera, GPS, bússola, acelerômetro, Bluetooth, EDGE, 3G e WiFi. A plataforma apresenta suporte para mídias de áudio, vídeo e imagem, nos formatos MPEG4, H.264, MP3, AAC, AMR, JPG, PNG, GIF, bem como acelerador gráfico 3D, baseados no OpenGL ES. Os dados podem ser armazenados em SQLite e a plataforma traz um navegador integrado com base no código livre do motor WebKit [Kanno, 2011].

4. Resultados e Discussões

Em virtude dos resultados parciais obtidos, compreende-se que a RVA pode auxiliar e ser utilizada no ensino da anatomia humana, tanto para estudantes de medicina e cursos afins quanto para os pacientes que frequentam os consultórios médicos.

O aplicativo para estudo da anatomia humana, utilizando RVA em Dispositivos Móveis, irá contribuir na formação dos estudantes. Foi verificada a existência de projetos similares, porém, segundo pesquisas realizadas, ainda não foi implementado algum trabalho que utilize RVA para dispositivos móveis, que é um recurso tecnológico atual, prático, moderno e interativo, contrapondo com os métodos de ensino tradicionais.

Em outras áreas, a utilização de RVA para dispositivos móveis está sendo inserida cada vez mais no cotidiano das pessoas, auxiliando as tarefas diárias devido à facilidade de manuseio, transporte e acessibilidade, como, por exemplo, os aplicativos LBS (Serviços Baseados em Localização), a navegação e mapas, a pesquisa visual, o reconhecimento facial e os controles de voz e códigos de barras. Logo, também é viável e eficaz o uso de dispositivos móveis com RVA a fim de promover um ensino com melhor acessibilidade.

Nesta aplicação proposta e através dos resultados e testes preliminares junto ao protótipo desenvolvido, o estudante poderá ter acesso em seu próprio dispositivo móvel adequado para a aplicação da RA, que exibirá as ilustrações como se fossem reais. Ao instalar o aplicativo em um celular que tenha sistema operacional Android, o usuário se deparará com uma tela inicial que traz algumas informações sobre os Sistemas do Corpo Humano, como visto na Figura 02. A partir daí, o usuário tem como opção aprender mais detalhadamente sobre suas divisões: o Sistema Nervoso Central - SNC (Figura 03), que é subdividido em encéfalo (Figura 03) e medula espinhal (Figura 04), e o Sistema Nervoso Periférico – SNP, que é subdividido em nervos (Figura 05) e gânglios (Figura 06). Além disso, ao clicar sobre o botão “Exibir em RA” de qualquer uma das telas, a imagem escolhida será exibida por meio da RA ao aproximar a câmera do dispositivo móvel sobre o marcador Android (Figura 01). Assim, a parte do corpo humano selecionada será exibida em RA, proporcionando aprendizagem mais interativa e interessante sobre a biologia humana.

Os marcadores são muito utilizados em RA para monitorar e verificar orientações dos objetos no ambiente, onde é possível definir previamente diversas destas orientações e para cada uma, especificar um comando concernente [Camilo-Junior, 2010]. Neste trabalho, foi utilizado um marcador já bastante popular para a programação em Android. Contudo, há maneiras de criar marcadores personalizados.

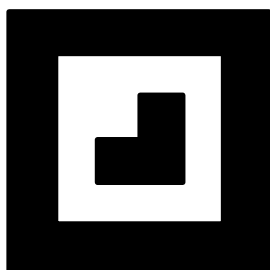


Figura 01: Marcador Android.

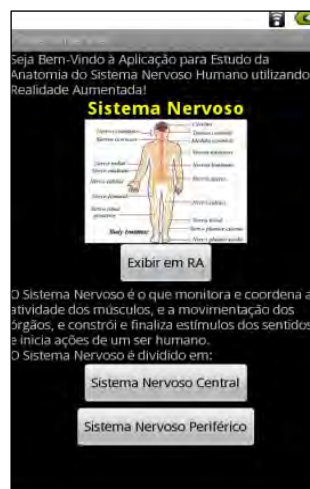


Figura 02: Sistema Nervoso no dispositivo móvel.



Figura 03: Encéfalo no dispositivo móvel.



Figura 04: Medula espinhal no dispositivo móvel



Figura 05: Nervos no dispositivo móvel.



Figura 06: Gânglios no dispositivo móvel.

Na Figura 07, por exemplo, o cérebro humano está sendo visualizado em Realidade Aumentada, por meio da câmara de um dispositivo móvel, como explicado anteriormente, o que possibilita um aprendizado mais dinâmico, visto que as imagens

em três dimensões provocam a imersão do usuário no sistema e também na própria anatomia do corpo humano.

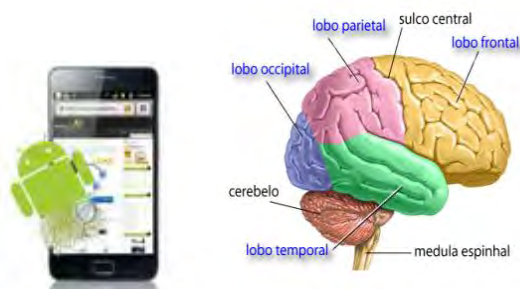


Figura 07: Cérebro humano em RA.

4. Conclusões

A evolução constante da tecnologia está impulsionando a educação para novos rumos, enfatizando a utilização de novas ferramentas e propiciando melhorias eficazes no processo de ensino/aprendizagem devido ao maior interesse e à motivação dos alunos. A medicina é uma das áreas de grande demanda ao uso de RVA em educação, considerando treinamento, diagnóstico, tratamento e simulação de cirurgias. Pelas suas características de visualização 3D e de interação em tempo real, permite a realização de aplicações médicas inovadoras, que antes não podiam ser realizadas.

A RVA é uma área tipicamente multidisciplinar que envolve conceitos provenientes de diversos segmentos, com aplicação em várias áreas ou campos específicos da vida econômica, social e cultural. Por permitir uma interação humano-computador mais natural em ambiente tridimensional (3D) e possibilitar a reprodução de situações reais, torna-se um recurso de amplo potencial [Costa e Ribeiro, 2009].

Acredita-se na importância dos professores trabalharem com as novas tecnologias de RVA, pois proporcionam a visualização e interação do estudante com o conhecimento abstrato de maneira completa, facilitando o caminho para a compreensão de teorias e conceitos.

A partir de pesquisas realizadas e análises conduzidas, verifica-se, que apesar de inúmeros sistemas desenvolvidos, ainda há vários desafios a serem superados a fim de que as aplicações de RVA sejam incluídas na rotina dos profissionais da área de saúde. Se por um lado tais desafios retardam o uso efetivo desta tecnologia, na prática, constituem oportunidades ímpares de pesquisa e desenvolvimento.

É fato que a Universidade possui o desafio de inserir no mercado de trabalho profissionais qualificados, com perfil crítico e investigador frente às diversas situações do cotidiano, com bom domínio da tecnologia, capaz de lidar com o pluralismo de dificuldades que podem ser encontradas em determinada população, estimulando o indivíduo a adquirir conhecimento científico e a desenvolver atitudes de cidadania em seu contexto social.

Perante o avanço tecnológico, já disponível em considerável parcela das instituições de ensino no Brasil, acredita-se haver possibilidades de mudanças das práticas pedagógicas em sala de aula, visto que a tecnologia já está inserida no dia-a-dia dos estudantes e que o processo ensino-aprendizagem deve ser condizente com a realidade que o aluno vivencia no seu cotidiano.

Referências

- Bodoff, S. et al.. Tutorial do J2EE. Rio de Janeiro, Editora Ciência Moderna Ltda., 2005.
- Bobbch G, et al.. UML: Guia do Usuário. Campus, 2005.
- Camilo-Junior, C. G.; Ueda, M. T. M.; Viana, R. F.. Um sistema de auxílio ao diagnóstico da escoliose baseado em RA. Revista Brasileira de Engenharia Biomédica, v. 26, n. 3, p. 185-193, dez. 2010.
- Cardoso, A. et al. Tecnologias para o desenvolvimento de sistemas de realidade virtual e aumentada. Recife: Universidade da UFPE, 2007. 222 p.
- Costa, R. M. e Ribeiro, M. W. Aplicações de realidade virtual e aumentada. Porto Alegre: SBC, 2009. 146 p.
- Chopard R P, Bonilauri A R C. O ensino da anatomia: uma visão discente. Rev Hosp Clin Fac Med Univ São Paulo 1997; 52(6): 337- 341.
- Drab, S. & Artner, N M. Motion Detection as Interaction Technique for Games & Applications on Mobile Devices. Pervasive Mobile Interaction Devices (PERMID 2005), Munich, Alemanha, 2005.
- Espírito S. A. M. *et al.* Uso de cadáveres no estudo de anatomia humana nas escolas da área da saúde. Ver Goiana Med 1981; 27(1/2): 107-116.
- Edelweiss M. I. Importância do estudo de necropsia (ou de peças cirúrgicas não fixadas) no ensino da Anatomia Patológica Macroscópica. Rev HCPA & Fac Med Univ Fed.Rio Gd do Sul 1993; 13(3).
- Harma, A. et al. Techniques and applications of wearable augmented reality audio. In: Audio Engineering Society Convention Paper, Amsterdam, Holanda, 2003.
- Kirner, C. e Siscoutto, R. Realidade virtual e aumentada: conceitos, projeto e aplicações. Porto Alegre: SBC, 2007. 202 p.
- Lamounier, E. e Cardoso, A. Realidade virtual: uma abordagem prática. São Paulo: Mania de Livro, 2004. 326 p.
- Lévy, P. As Tecnologias da Inteligência: O Futuro do Pensamento na Era da Informática. Trad. Carlos Irineu da Costa. Rio de Janeiro: Ed. 34, 1993. 208p.
- Martins, A. S. F. Tecnologia Educacional e os Recursos Pedagógicos. In: Anuário da Produção Acadêmica Docente. Vol. III, n6, 2009.
- Neto, S. et al. Criação de Aplicações de Realidade Aumentada em Dispositivos Móveis Baseados em Symbian OS. In: II Workshop de Aplicações de Realidade Virtual, Universidade Federal de Pernambuco, 2006, p. 16-19.
- Quartiero, E. M. As Tecnologias da Informação e Comunicação e a Educação. Revista Brasileira de Informática na Educação, vol. 4, 1999.
- Reimann, C. Kick-real, a Mobile Mixed Reality Game. In: Proceedings of the International Conference on Advances in Computer Entertainment Technology. Valencia, 2005, p. 387.
- Siscoutto, Robson; Costa, Rosa Maria. Realidade virtual e aumentada: uma abordagem tecnológica. Porto Alegre: SBC, 2008. 357 p.

Aplicação de ferramentas automatizadas para geração de formulários Web acessíveis

Fernanda B. Faria¹, Thiago J. Bittar^{1,2}, Luanna L. Lobato¹

¹Departamento de Ciências da Computação – Universidade Federal de Goiás (UFG)
CEP: 75704-020 – Catalão – GO – Brasil

² Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP)
Caixa Postal 668 – 13560-970 – São Carlos – SP – Brasil

ferrbontempo@gmail.com, jabur@icmc.usp.br, lll@catalao.ufg.br

Abstract. *This paper shows how the use of automated tools may help in the generation of accessible forms. Therefore, we used two tools that assist the developer in this task, a form to generate itself, called Quick Form Builder and another to validate the code generated by this tool, the W3C code validator. Thus it is apparent that the tools help developers to develop affordable forms, since it does not need to meet accessibility guidelines, simply use the tools that are intuitive and easy to access.*

Resumo. *Neste trabalho é apresentado como a utilização de ferramentas automatizadas podem ajudar na geração de formulários acessíveis. Para tanto, foram utilizadas duas ferramentas que auxiliam ao desenvolvedor nesta tarefa, uma para gerar formulário em si, chamada Quick Form Builder e outra para validar o código html de formulários Web, o validador de código do W3C. Com isso, é perceptível que as ferramentas auxiliam o desenvolvedor a desenvolver formulários acessíveis, visto que o mesmo não precisa conhecer as diretrizes de acessibilidade, é necessário apenas usar as ferramentas que são intuitivas e de fácil acesso.*

1. Introdução

Para que haja um desenvolvimento ágil e com mais qualidade é necessário que se tenha ferramentas que agregam mais facilidades e recursos aos desenvolvedores e que evoluam de acordo com suas necessidades. Uma preocupação maior que se tem, além do desenvolvimento rápido, é que sistemas Web possuam acessibilidade, para que a inclusão digital seja de fato efetivada. Assim, é necessário produzir artefatos Web sem barreiras, ou seja, eliminar as dificuldades de acesso às pessoas com deficiência e para isso, podem ser utilizadas ferramentas para auxiliar o desenvolvedor e também para que este tenha conhecimento das diretrizes de acessibilidade e como aplicá-las [Sloan et al. 2006], [Maia 2010].

Acessibilidade Web está relacionada com a possibilidade de que diferentes perfis de usuários e, principalmente, pessoas com deficiência sejam capazes de usá-la [Lima 2007]. Mais precisamente, uma Web com conteúdo interativo, fácil de entender, compreender e manipular por pessoas com os mais diversos tipos de deficiência [W3C 2008].

Para interagir em sistemas e sites Web, na maioria das vezes é necessário o preenchimento de formulários, seja para cadastro ou simplesmente para receber uma informação de retorno. Entretanto, como visto em um estudo de caso recente [Bittar et al. 2012a], os formulários, sobretudo das universidades públicas do Brasil, não atendem às diretrizes de acessibilidade, o que dificulta o acesso pelos usuários, principalmente aqueles que possuem algum tipo de deficiência.

Os problemas identificados nos formulários Web pelos autores supracitados, ressaltam a importância de ferramentas automatizadas para auxiliar o desenvolvedor no desenvolvimento de tais recursos. Assim, é visto no presente trabalho como ferramentas automatizadas auxiliam na geração de formulários acessíveis. Para tanto, é mostrado o uso de um gerador de formulários acessíveis, *Quick Form Builder*¹, desenvolvido pela empresa *Accessify*², que foi escolhido devido a fácil utilização e popularização da ferramenta, juntamente com um validador de código, do *World Wide Web Consortium* (W3C)³, para que se comprove a eficácia do gerador de formulários.

O trabalho é apresentado da seguinte forma: Na Seção 2 é mostrada a metodologia utilizada para a realização da pesquisa. Na Seção 3 tem-se os trabalhos relacionados. Na Seção 4 é visto o funcionamento do gerador de formulários acessíveis, em que seu código é testado pelo validador na Seção 5, na qual mostra como o mesmo funciona. Na Seção 6 são mostradas as contribuições que são obtidas utilizando essas ferramentas. E, por fim, na Seção 7, tem-se a conclusão do que foi feito.

2. Metodologia

Inicialmente, foi feito um estudo envolvendo as diretrizes de acessibilidade, tanto para conteúdo *Web Content Accessibility Guidelines 2.0* (WCAG 2.0) [W3C 2008], quanto para ferramentas de autoria *Authoring Tool Accessibility Guidelines* (ATAG) [W3C 2000].

Posteriormente, verificou-se a acessibilidade em formulários de contato em sites de 109 universidades públicas brasileiras, procurando abranger o máximo possível deste cenário, utilizando diretrizes do WCAG 2.0, na qual constatou-se falta de acessibilidade nos critérios analisados: formulários feitos utilizando tabelas, o que não é recomendando; formulários feitos com *labels*; formulários que utilizam CAPTCHA (que é uma ferramenta contra acessos de robôs ou mecanismos que não sejam pessoas); formulários que validam campos; e formulários que funcionam corretamente [Bittar et al. 2012a]. Após essa verificação, foi feita a análise em ferramentas de autoria, utilizando as diretrizes da ATAG, e foi constatado que nenhuma delas oferece o suporte adequado para desenvolvimento de conteúdo acessível [Bittar et al. 2012b].

Visto este cenário e o estudo realizado por Trewin *et al.* (2010), que mostrou que a maioria dos desenvolvedores não utilizam recursos e diretrizes de acessibilidade no desenvolvimento por falta de conhecimento dos mesmos, faz-se necessário o uso de ferramentas que auxiliem aos desenvolvedores a suprir essa deficiência.

Logo, é feito neste trabalho uma análise de um gerador de formulários acessíveis automatizado, *Quick Form Builder*, mostrando cada etapa para o desenvolvimento dos

¹<http://accessify.com/tools-and-wizards/accessibility-tools/quick-form-builder/>

²<http://accessify.com/about/>

³<http://validator.w3.org/>

formulários, de maneira intuitiva e testando o formulário gerado no validador do W3C. Com isso, pode-se analisar como essas ferramentas podem contribuir no processo de desenvolvimento acessível.

3. Trabalhos relacionados

Bittar *et al.* (2012a) mostram a verificação de formulários de contato em universidades públicas brasileiras e constata a defasagem do uso das diretrizes do WCAG 2.0, visto que haviam formulários feitos com tabelas, outros que utilizam CAPTCHA sem recurso auditivo, o que prejudica o acesso de pessoas com deficiência à informação. Visto que, a leitura do CAPTCHA é feita independentemente da leitura do formulário, pois leitores de tela leem apenas o conteúdo do formulário e é um robô realizando esta leitura, o CAPTCHA precisa prover recursos para ser ouvido.

Bittar *et al.* (2012b) apresentam um estudo sobre as principais ferramentas de autoria utilizadas por desenvolvedores no Brasil tais como: *Dreamweaver*®, Eclipse (Helios)®, *Netbeans 7.1*®, NVU 1.0®, e *Microsoft Expression Web*®, foi constatado que nenhuma dessas ferramentas segue corretamente todos os critérios de acessibilidade analisados, o que dificulta o desenvolvimento de sistemas Web acessíveis.

Além de acessibilidade em formulários, existem outros trabalhos, como o apresentado por Santos (2012), que trata da acessibilidade em menus na Web para pessoas de meia idade, verificando as dificuldades enfrentadas por aqueles que possuem problemas de visão. O que também poderia ser evitado com o conhecimento das diretrizes de acessibilidade ou com ferramentas que auxiliassem aos desenvolvedores a seguir esses critérios.

Já em Sousa (2011) é feita uma avaliação em portal de instituições de ensino superior tendo em foco pessoas com deficiência visual, analisando portais da Universidade de Pernambuco (UPE), Universidade Federal do Rio de Janeiro (UFRJ) e Universidade Federal do Estado do Rio de Janeiro (UNIRIO), nos quais também são encontrados problemas de acessibilidade, que poderiam ser evitados.

4. Ferramenta de geração de formulários acessíveis

Para auxiliar no desenvolvimento de formulários acessíveis existem algumas ferramentas disponíveis ao desenvolvedor de rápida e intuitiva manipulação. Ou seja, ferramentas que geram formulários sem utilização de tabelas, utilizando *labels* de maneira adequada, de acordo com o critério 3.3.2 do WCAG 2.0 consistente de que quando o conteúdo exigir a entrada de dados pelo usuário é obrigatório o uso de rótulos, para evitar erros [W3C 2008]. Verificam se o fechamento as *tags* está sendo feito corretamente, o que condiz com o critério 4.1.1 do WCAG 2.0, na qual consiste que os elementos que dispõem de marcas de início e de fim devem ser completos, ou seja, devem ser abertos e fechados corretamente, caso contrário são considerados atributos incorretos ou incompletos [W3C 2008].

Como exemplo, tem-se o *Quick Form Builder* que tem por objetivo gerar conteúdo acessível e também ser de fácil manuseio ao desenvolvedor, o que é extremamente necessário visto que o operador desta ferramenta é o desenvolvedor de sistemas para Web que utilizam formulários de contato.

Na Figura 1, é ilustrada a página principal da ferramenta, em que devem ser definidos campos que o formulário precisa ter. Caso o usuário não saiba o que inserir, existem

alguns exemplos de dados que podem ser inseridos e, também, logo abaixo ao local de inserção, tem-se as instruções que se devem ser seguidas para usar a ferramenta adequadamente.

Quick Form Builder

Step 1 - list all form fields

Try it out! Type the field names out first, e.g.:

```
First name
Surname
Town
State/Region
```

[Next step \(define form control types\) →](#)

Figura 1. Página inicial da ferramenta, primeiro passo

Após a conclusão da etapa inicial, ilustrada na Figura 1, pressiona-se o botão *Next Step*, para que seja executado o próximo passo do desenvolvimento do formulário no qual se definirá os tipos destes campos, como pode ser visto na Figura 2.

Quick Form Builder

Step 2 - define form control type

☐ Mark all fields as required

Nome:	Text input	☐ Required field?
Email:	Text input	☐ Required field?
Assunto:	Text input	☐ Required field?
Mensagem:	Text input	☐ Required field?

[← Back to previous step](#) [Next step \(choose some markup options\) →](#)

Figura 2. Segundo passo, opções de marcação

Na Figura 2 é apresentado como se define cada tipo dos campos do formulário bem como os campos obrigatórios a serem preenchidos pelo usuário. Esta etapa é significativamente importante, pois com os campos obrigatórios torna-se mais fácil o preenchimento do formulário.

Na Figura 3 são apresentados os tipos de marcação, como de sintaxe, se será XHTML ou HTML. A forma de envio de dados podem ser por meio dos métodos post ou get. Além disso, o usuário pode definir o nome do formulário, a mensagem de erro caso os dados não sejam inseridos corretamente pelo usuário final e a mensagem do botão de submissão.

Quick Form Builder

Step 3 - choose some markup options

Markup syntax

☒ XHTML ☐ HTML

Form method

☒ post ☐ get

☒ Create `class/ids` based on label text (e.g. "Middle Name(s)" becomes "MiddleNames")

☒ Prefix field `ids` with "txt", "sel" etc (e.g. "Middle Name(s)" becomes "txtMiddleNames")

☒ Create ARIA `labelledby` attributes (what's your labelledby?)

Form `action`:

Form `name/id`:

Help text to display for required fields: (will appear in `required` attribute of form control)

Submit button value:

[← Back to previous step](#) [Next step \(generate the form\) →](#)

Figura 3. Terceiro passo, opções de marcação

Concluindo o passo 3, ocorre a geração do código fonte do formulário, que é ilustrado na Figura 4. Com a geração do código, está finalizado o desenvolvimento do formulário.



Figura 4. Quarto passo, geração de código fonte

Após visualizar o código, o desenvolvedor tem a opção de voltar ao passo anterior, ou começar um novo formulário, retornando ao passo 1.

O código gerado possui os *labels* especificados e validação dos campos, mas ao executar o código é perceptível que não há a marcação dos campos obrigatórios na interface, o que pode prejudicar na prevenção de erros pelo usuário, que é uma falha de acessibilidade. A ferramenta também não gera o recurso de CAPTCHA, o que torna o sistema menos seguro, porém mais acessível.

O formulário da Figura 5 possui os campos de nome, email, assunto e mensagem, que são os campos básicos para um formulário de contato, é uma tela simples e de fácil entendimento para o usuário.

Formulário de Contato

Nome:

Email:

Assunto:

Mensagem:

Enviar

Figura 5. Formulário gerado com a ferramenta Quick Form Builder

Na próxima seção é apresentado o validador de códigos do W3C, que faz a verificação do código gerado pela ferramenta *Quick Form Builder* aqui apresentada, analisando se o mesmo é usável ou não e apontando os erros cometidos.

5. Validador de código do W3C

O validador de código do W3C é um *software* livre utilizado para verificar o código fonte de páginas, apontando erros de grafia, uso impróprio de CSS e também alerta o desenvolvedor sobre potenciais erros de acessibilidade [W3C 2007].

A inserção do código em HTML pode ser feita por três vias, inserindo-se a URL da página que se deseja validar, fazendo *upload* de arquivos locais ou por entrada direta, na qual se coloca o código fonte diretamente no validador, como pode ser observado na Figura 6. Cada aba do validador indica uma forma de entrada do código a ser analisado, é

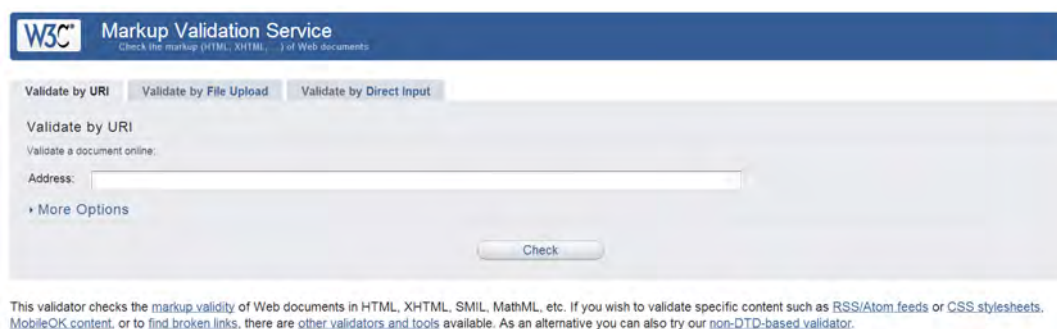


Figura 6. Interface do validador de código do W3C

um sistema simples e intuitivo para os desenvolvedores, que desejam validar seus códigos em HTML.

Na Figura 7 é ilustrada a inserção de código (que foi realizada neste trabalho) e faz-se a validação do mesmo. Neste caso, o mesmo foi verificado com sucesso, ou seja, não possui erros de formação, o que é importante para prover acessibilidade.

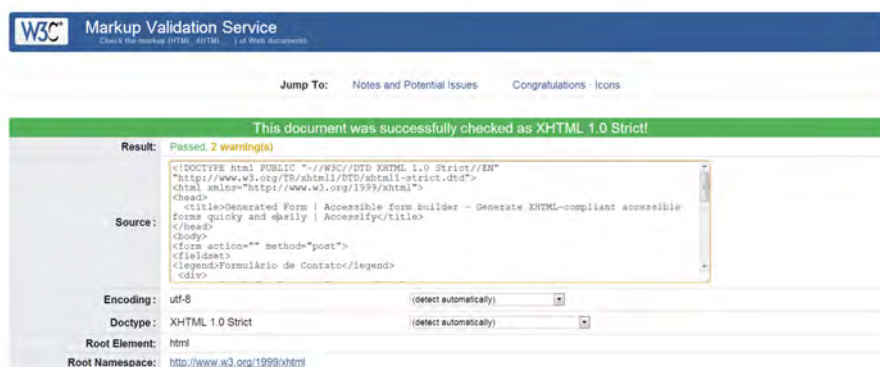


Figura 7. Validando o código

Caso o arquivo possua erros ou *warnings*, o sistema apresenta uma mensagem ao desenvolvedor informando-o sobre os erros e suas respectivas localizações no código, como mostrado nas Figuras 8 e 9, nas quais foram utilizados o mesmo código da Seção 4 removendo-se os cabeçalhos do mesmo.

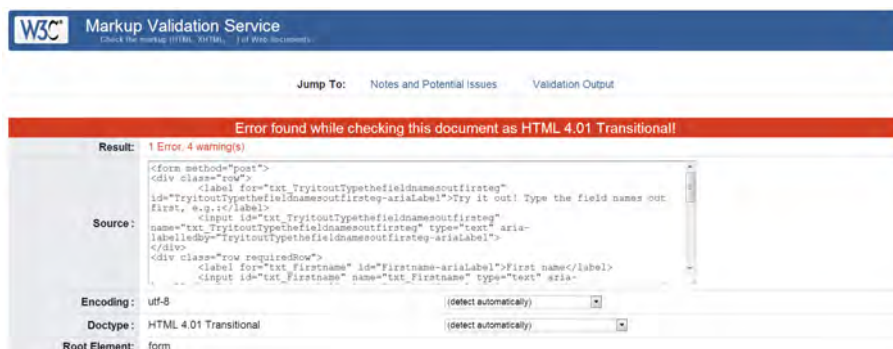


Figura 8. Código com erro e *warning*

Na Figura 8 é visto que 1 erro foi encontrado no documento e 4 *warnings*, logo, o código não passou na avaliação validador. Quando isso ocorre, o desenvolvedor pode adequar seu código no próprio editor da ferramenta e revalidá-lo, ou resubmetê-lo sempre que necessário.

Já na Figura 9 é mostrado como o validador indica ao usuário os erros encontrados e sugestões de correção, o que torna mais fácil a adequação do código, visto que o desenvolvedor pode não ter conhecimento do que fez de errado e nem onde cometeu tal erro.



Figura 9. Indicando quais erros encontrados

6. Contribuições da utilização das ferramentas

Com a utilização destas ferramentas torna-se mais fácil ao desenvolvedor prover acessibilidade em formulários, pois pode-se validar o código verificando se está bem formado. Além disso, caso hajam erros, o validador avisa qual tipo e em qual parte do código está, o que auxilia na identificação dos mesmos e na sua correção efetiva.

Assim, com esse auxílio, torna-se mais fácil o desenvolvimento de formulários acessíveis, visto que não é necessário conhecer as diretrizes de acessibilidade pois o validador auxilia nesta parte.

7. Conclusão

Foi mostrado neste trabalho uma ferramenta que auxilia o desenvolvimento de um formulário acessível, sendo a mesma intuitiva e de fácil manuseio. Para tanto, foi realizado teste em um validador de código para inspecionar se o formulário é realmente acessível. Percebeu-se então que, o código gerado pelo gerador de formulário é acessível e usável. Contudo, verifica-se que a ferramenta não indica quais são os campos obrigatórios o que também não foi condenado pelo validador de código. Com isso, pode-se sugerir melhorias para a ferramenta de autoria tais como, mostrar os campos obrigatórios do formulário.

Logo, percebe-se que os desenvolvedores possuem uma alternativa para desenvolver formulários acessíveis, visto que ambas as ferramentas são *software* livre, ou seja, não gera custo adicional e que são de fácil acesso.

É proposto como trabalho futuro a realização de um estudo de caso para melhorar a ferramenta, atrelando-a a um validador de acessibilidade, para que em uma única ferramenta se tenha as correções efetivadas de maneira imediata.

Referências

- Bittar, T., Faria, F., do Amaral, L., and de Mattos Fortes, R. (2012a). An assessment of accessibility in contact forms of brazilian public universities. In *Information Systems and Technologies (CISTI), 2012 7th Iberian Conference on*, pages 1 – 6, Madrid, Espanha. IEEE Xplorer.
- Bittar, T. J., do Amaral, L. A., Faria, F. B., and de Mattos Fortes, R. P. (2012b). Supporting the developer in an accessible edition of web communications: a study of five desktop tools. In *Proceedings of the Workshop on Information Systems and Design of Communication*, ISDOC '12, pages 3–9, New York, NY, USA. ACM.
- Lima, S. T. (2007). Avaliação da acessibilidade de sítios web por meio de m'etricas de software. Mestrado em gestão de conhecimento e da tecnologia da informação, Universidade Católica de Brasília, Brasília - DF.
- Maia, L. S. (2010). Um processo para o desenvolvimento de aplicações web acessíveis. Mestrado em ciências da computação, Universidade Federal de Mato Grosso do Sul.
- Santos, E. P. B. (2012). Acessibilidade em menus de navegação horizontais na web para pessoas de meia-idade. Mestrado em ciências de computação e matem'atica computacional, Universade São Paulo, São Carlos - SP.
- Sloan, D., Heath, A., Hamilton, F., Kelly, B., Petrie, H., and Phipps, L. (2006). Contextual web accessibility - maximizing the benefit of accessibility guidelines. In *Proceedings of the 2006 international cross-disciplinary workshop on Web accessibility (W4A): Building the mobile web: rediscovering accessibility?*, W4A '06, pages 121–131, New York, NY, USA. ACM.
- Sousa, M. F. C. (2011). Acessibilidade web: Uma avaliação em portal de instituições de ensino superior visando pessoas com deficiência visual. Monografia, Universidade de Pernambuco, Pernambuco. p. 89.
- Trewin, S., Cragun, B., Swart, C., Brezin, J., and Richards, J. (2010). Accessibility challenges and tool features: an ibm web developer perspective. In *Proceedings of the 2010 International Cross Disciplinary Conference on Web Accessibility (W4A)*, W4A '10, pages 32:1–32:10, New York, NY, USA. ACM.
- W3C (2000). Authoring tool accessibility guidelines 2.0. <http://www.w3.org/WAI/intro/atag.php>. Acessado em 10 de fevereiro de 2012.
- W3C (2007). Serviço de validação de css. <http://jigsaw.w3.org/css-validator/about.html.pt-BR>. Acesso em 25 de janeiro de 2013.
- W3C (2008). Web content accessibility guidelines (wcag) 2.0. <http://www.w3.org/TR/WCAG20/>. Acessado em 10 de fevereiro de 2012.

Otimização das Bandas Fotônicas Proibidas Absolutas de Cristais Fotônicos Anisotrópicos Bidimensionais utilizando Sistemas Imunológicos Artificiais

Anderson Dourado Sisnando¹,
Vitaly Felix Rodriguez Esquerre¹, Francismari Noronha dos Santos^{1,2}

¹Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal da Bahia – Salvador, BA – Brasil.

²Departamento de Tecnologia Eletro-Eletrônica, Instituto Federal de Educação, Ciência e Tecnologia (IFBA) – Salvador, BA – Brasil.

anderson.sisnando@ufba.br, vitaly.esquerre@ufba.br,
mari.noronhas@gmail.com

Abstract. *The finite element method (FEM) and artificial immune system (AIS) were used to solve the inverse problem involving the search and optimization of photonic bandgap (PBG) absolutely crystal structures with periodicity in two dimensions. Square and triangular arrays composed of tellurium (anisotropic material) and air was analyzed. The PBGs were calculated along the first Brillouin zone while avoiding the high computational effort in this kind of task, obtaining the midgap ratio absolute considering PBGs between modes TE_{12} and TM_{34} , of 21.27% and 28.45% for the square and triangular arrays, respectively.*

Resumo. *O método dos elementos finitos (FEM) e o sistema imunológico artificial (SIA) foram utilizados para resolver o problema inverso envolvendo a busca e otimização da banda fotônica proibida (PBG) absoluta de estruturas cristalinas com periodicidade em duas dimensões. Arranjos quadrados e triangulares compostos por telúrio (Material Anisotrópico) e de ar foram analisados. As PBGs foram calculadas ao longo da primeira região de Brillouin, evitando o alto esforço computacional neste tipo de tarefa, obtendo-se a relação da banda absoluta com seu valor médio, considerando PBGs entre os modos TE_{12} e TM_{34} , de 21,27% e 28,45% para os arranjos quadrados e triangulares, respectivamente.*

1. Introdução

Dispositivos baseados em cristais fotônicos (CF) são de grande aplicação nos sistemas de comunicações ópticas, tais como: demultiplexadores, cavidades ressonantes, filtros entre outros [Joannopoulos, et al 2008]. Uma forma de otimizar esses cristais implica em obter estruturas que apresentem uma maior banda fotônica proibida (PBG) para os modos de propagação desejados, o que permite uma maior largura de banda de operação. A PBG absoluta em um CF ocorre quando ambos os modos de polarização exibem superposição de suas bandas fotônicas proibidas (ou seja, a polarização independente). Em qualquer CF a busca e otimização de PBGs absolutas são obtidas através da exploração dos parâmetros: tipo de estrutura, fator de preenchimento, os índices de refração dos materiais, forma e através da redução de simetria.

1.1. Trabalhos Relacionados

Este tipo de problema é abordado em [Malheiros, Esquerre e Figueroa 2011], [Li, Gu e Yang 1998] e [Li, Gu e Yang 1999] com arranjos quadrados e triangulares e PBGs absolutas foram obtidas. Os efeitos isolados e combinados da redução de simetria e a introdução orifícios de ar nos elementos de telúrio do cristal foram estudados em [Zaccaria, et al 2008].

A maximização da PBG absoluta é explorada em [Rezaei, et al 2009], com cristais fotônicos [Kurt e Citrin 2005], usando também telúrio em diferentes arranjos, considerando diversas geometrias no interior dos cristais (circular, triangular, elíptica, retangular e quadrada). Em todos os casos, o método de expansão de ondas planas (PCE) é utilizado. No entanto, a propagação da luz no plano de CFs com arranjos 2-D, é investigada em [Rodriguez-Esquerre, et al 2006] e [Lin e Jao 2007], usando o FEM no domínio da frequência e os resultados encontrados mostram concordância quando comparados com os resultados obtidos pelo método PCE. O FEM é uma representação numérica muito flexível e uma eficiente técnica de modelagem de dispositivos/componentes não homogêneos e de estruturas complexas. Além disso, em comparação com o método das diferenças finitas no domínio do tempo (FDTD) exibe convergência da mesma ordem e não tem limitações de modelagem para estruturas não homogêneas.

Neste trabalho, utiliza-se CF compostos por telúrio e ar em ambos os arranjos (quadrado e triangular) considerando o problema inverso para os modos transversais magnéticos e elétricos (TE_{12} e TM_{34}) através do SIA, com o objetivo de comparar com os resultados obtidos em [Malheiros, Esquerre e Figueroa 2011], que até o momento obteve melhores resultados, utilizando como método o algoritmo genético (AG). Uma característica importante de utilizar o SIA ao invés do AG é a possibilidade de achar novos ótimos, já que esse tipo de algoritmo poderá aumentar o espaço de busca, por possibilitar através da clonagem e afinidade uma maior diversidade na população [Castro 2006]. O mesmo é utilizado em problemas eletromagnéticos em [Silva-Santos, et al 2010a], [Silva-Santos, et al 2010b] e [Dourado-Sisnando, et al 2012].

Para o cálculo das PBGs, uma formulação baseada no FEM [Rodriguez-Esquerre, et al 2006], [Quinónez, et al 2006] e [Shi 2010] é utilizada. Embora a estratégia proposta é aplicada para obter PBGs absolutas em materiais anisotrópicos, a mesma também pode ser aplicada para problemas mais simples, tais como o caso isotrópico e também para uma única polarização [Dourado-Sisnando, et al 2012].

Na próxima seção, uma introdução sobre cristais fotônicos (geometria, problema abordado e arranjos) é apresentada. A seção 3 aborda a aplicação do SIA em conjunto com o FEM, logo depois, os resultados numéricos e as geometrias otimizadas são discutidas e comparadas.

2. Cristais Fotônicos Bidimensionais

Num cristal fotônico, a luz que atravessa o mesmo é espalhada pela geometria periódica, sendo refratada e parcialmente refletida nas interfaces entre os materiais. O padrão resultante das ondas será reforçado ou cancelado dependendo do comprimento de onda da luz, os índices de refração dos materiais, sua direção de propagação através do cristal, o tamanho e arranjo das estruturas (geometria), e a partir dessas iterações aparece à banda fotônica proibida [Malheiros, Esquerre e Figueroa 2011], [Dourado-Sisnando, et al 2012].

Por apresentarem simetrias, os cristais, podem ter a análise dos campos e modos eletromagnéticos a partir de células unitárias com condições de contorno periódicas. A partir desta célula é obtida a primeira zona de *Brillouin*, a qual define a região do espaço dos vetores de onda que determinam os modos de propagação, Figura 1.

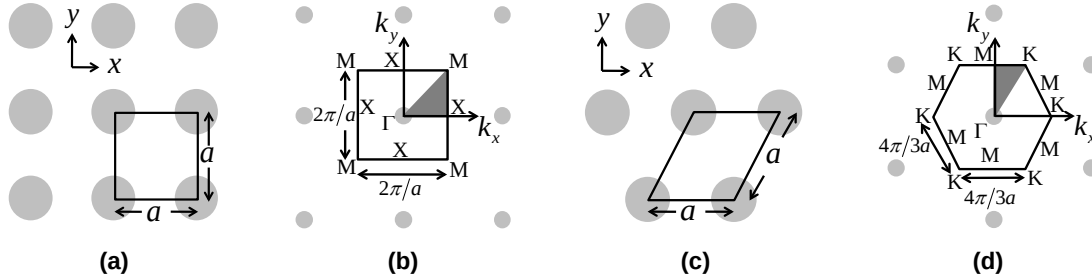


Figura 1. (a) Célula unitária do arranjo quadrado. (b) região de Brillouin. (c) célula unitária do arranjo triangular. (d) região de Brillouin, [Malheiros, Esquerre e Figueroa 2011].

Devido à periodicidade da estrutura cristalina, discretiza-se apenas uma célula unitária e aplicam-se condições de contorno periódicas para tornar iguais os campos do topo e da parte inferior, e os campos dos lados esquerdo e direito, da célula unitária. Os valores de k_x e k_y são restritos a primeira região de *Brillouin*, Figura 1.

Os dois arranjos utilizados, Figura 1, são representados por elementos circulares no interior da célula unitária apenas para simplificar o entendimento. A análise do FEM em cristais fotônicos 2-D pode ser tratada separadamente para os modos TM e TE como em [Malheiros, Esquerre e Figueroa 2011].

A região sombreada (triângulo) nas Figuras 1(b) e 1(e) correspondem à região irreduzível de Brillouin limitadas por $\Gamma=(0,0)\pi/a$, $X=(1,0)\pi/a$, $M=(1,1)\pi/a$ para o arranjo quadrado e por $\Gamma=(0,0)\pi/a$, $K=(2/3, 2\sqrt{3}/3)\pi/a$, e $M=(0, 2\sqrt{3}/3)\pi/a$ para o arranjo triangular.

Neste trabalho, a configuração interna da célula unitária (disposição do material) é considerada, como apresentando a geometria mais arbitrária. Em tal caso, é necessário resolver em todos os pontos marcados porque a máxima e mínima frequências estão naqueles pontos. Logo, calcula-se a PBG para nove e treze pontos, arranjos quadrados ou triangulares, respectivamente.

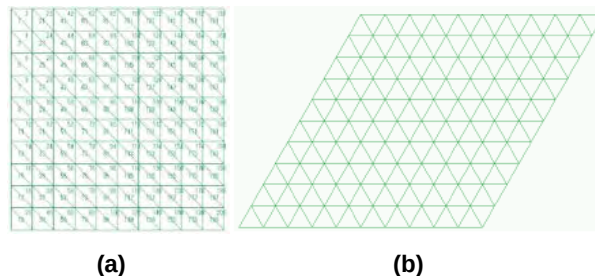


Figura 2. Célula unitária dos cristais fotônicos dividida em 200 triângulos correspondente ao arranjo (a) quadrado e (b) triangular, respectivamente [Malheiros, Esquerre e Figueroa 2011].

As PBGs absolutas são obtidas, com o anisotrópico uniaxial (telúrio) o qual possui dois valores de índice de refração: $n_e = 6,2$ (Extraordinário) e $n_o = 4,8$ (ordinária) para comprimentos de onda entre $3,5\mu\text{m}$ e $35\mu\text{m}$. O eixo extraordinário é considerado

ao longo do eixo z [Li, Gu e Yang 1998] e [Li, Gu e Yang 1999]. As malhas correspondentes às células unitárias são apresentadas na Figura 2. Cada triângulo representa uma parte dos anticorpos e ela poderá ser de ar ou telúrio.

3. Sistema Imunológico Artificial

O SIA é utilizado no intuito de otimizar as bandas fotônicas proibidas absolutas de um cristal composto de telúrio e ar, com arranjos quadrados e triangulares para os modos de polarização TE_{12} e TM_{34} simultaneamente. A banda fotônica proibida é calculada através dos nós da primeira zona de *Brillouin* [Li, Gu e Yang 1998].

Para todas as simulações são consideradas uma população de 20 e 50 indivíduos (anticorpos), com evolução acontecendo durante 1000 gerações. A otimização da banda fotônica proibida utilizando os algoritmos bio-inspirados, se mostram uma boa opção, pois o cálculo das bandas proibidas requer tempo de computação relativamente longo, o que torna inviável realizar uma varredura em todo o universo de possibilidades, que corresponde a 2^{200} . Deste modo, o SIA é escolhido, assim como outros algoritmos evolutivos foram anteriormente utilizados [Malheiros, Esquerre e Figueroa 2011], e [Li, Gu e Yang 1999] e [Zaccaria, et al 2008] para a solução de problemas eletromagnéticos.

O sistema imunológico artificial considerado possui as seguintes etapas:

a) Inicialização: A população inicial (anticorpos) é gerada e cada indivíduo é representado como uma sequência binária. Se uma posição é "1", o respectivo triângulo é composto por telúrio, caso contrário, o triângulo considerado será composto de ar. Considera-se uma célula unitária para cada malha, como observado na Figura 2. Elas foram divididas em 200 elementos triangulares, resultando em 441 nós, cada triângulo representa uma posição em cada indivíduo (anticorpo).

b) Avaliação: A aptidão (fitness) de cada indivíduo é avaliada pela função objetivo dada por:

$$\text{Função Objetivo} = \frac{\text{Frequência}_{\text{superior}} - \text{Frequência}_{\text{inferior}}}{\text{Frequência}_{\text{média}}} \quad (8)$$

c) Seleção: Os 50% melhores indivíduos (anticorpos) da população gerada são selecionados para clonagem e sua probabilidade de ser selecionado é proporcional à sua aptidão, a seleção por ranking, é considerada.

d) Clonagem: selecionados, os indivíduos são clonados, de modo a ter novamente o tamanho da população inicial, formando uma população composta dos melhores indivíduos.

e) Mutação: a mutação é realizada ocorre a cada geração.

g) Afinidade: os indivíduos que resultam da mutação são avaliados por uma afinidade (proximidade) de 0,02 (valor médio selecionado através da observação da evolução do algoritmo) para os indivíduos que foram selecionados na etapa de seleção antes da clonagem.

h) Avaliação: O *fitness* de cada indivíduo é avaliado pela função objetivo Eq. 8.

i) Elitismo e Supressão: é de grande importância, pois garante que os melhores indivíduos permanecem na próxima geração, evitando uma possível redução do fitness (aptidão), afetando o processo de convergência. Os três piores indivíduos da geração

atual são substituídos pelos três melhores da geração anterior.

j) Atualização: os indivíduos criados em cada geração são inseridos na população.

k) Fim: O critério de parada considerado é o número de gerações.

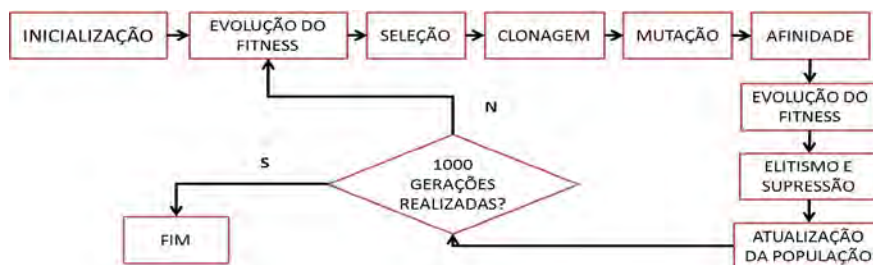


Figura 3. Fluxograma simplificado do SIA utilizado.

4. Resultados Numéricos e Discussão

A plataforma computacional utilizada foi um computador com processador Intel® Core™ i5 2310 (2.90 GHz, 6MB Cache L2), 8GB de RAM com o Windows 7 Home Basic. É considerada a existência de PBGs absolutas entre os modos TE 1-2 e TM 3-4. O cálculo da banda proibida é realizado em 9 e 13 pontos da primeira região de Brillouin para os arranjos quadrados e triangulares, respectivamente. Isto reduz significativamente o esforço computacional e o tempo de processamento. Os resultados obtidos para os arranjos quadrados e triangulares são apresentados nesta seção.

4.1. Arranjo Quadrado e Arranjo Triangular

As evoluções dos *fitness* em função do número de gerações, bem como as células unitárias resultantes são apresentadas na Figura 4 para o arranjo quadrado e arranjo triangular. Cada geração é composta por vinte (Figura 4 (a) e (c)) e cinquenta (Figura 4 (b) e (d)) indivíduos (anticorpos), nota-se uma evolução rápida do algoritmo como um todo. A diversidade de indivíduos com diferentes *fitness* é também observada, o que indica que houve uma exploração razoável de busca global caracterizada por indivíduos de baixa aptidão. As configurações das células unitárias obtidas para o arranjo quadrado e o arranjo triangular sobre uma malha de 10×10 triângulos são apresentadas como uma inserção na Figura 4 (a) - (d). Por outro lado, muitos indivíduos apresentam um bom valor de *fitness*. Isto pode ser atribuído a discretização realizada nas células unitárias para ambos os arranjos. Os cristais otimizados são apresentados na Figura 5. As áreas escuras são telúrio. Os diagramas de dispersão destas estruturas podem ser vistos na Figura 6.

A convergência dos resultados confirma que os cristais obtidos pela otimização com o SIA são melhores que em [Malheiros, Esquerre e Figueroa 2011]. As PBGs absolutas encontradas pelo SIA são maiores que as encontradas pelo AG em [Malheiros, Esquerre e Figueroa 2011], Tabela 1, demonstrando que o algoritmo desenvolvido é também uma ótima ferramenta para este tipo de problema.

As geometrias resultantes são consideradas como sendo regiões de ar em Telúrio [Malheiros, Esquerre e Figueroa 2011], ao contrário do encontrado em outros trabalhos [Li, Gu e Yang 1998] e [Li, Gu e Yang 1999], onde PBGs absolutas podem ser facilmente encontradas considerando geometrias compostas por hastes dielétricas de

telúrio no ar. É adotada como estratégia de refinamento a otimização de uma célula unitária dividida em 200 triângulos.

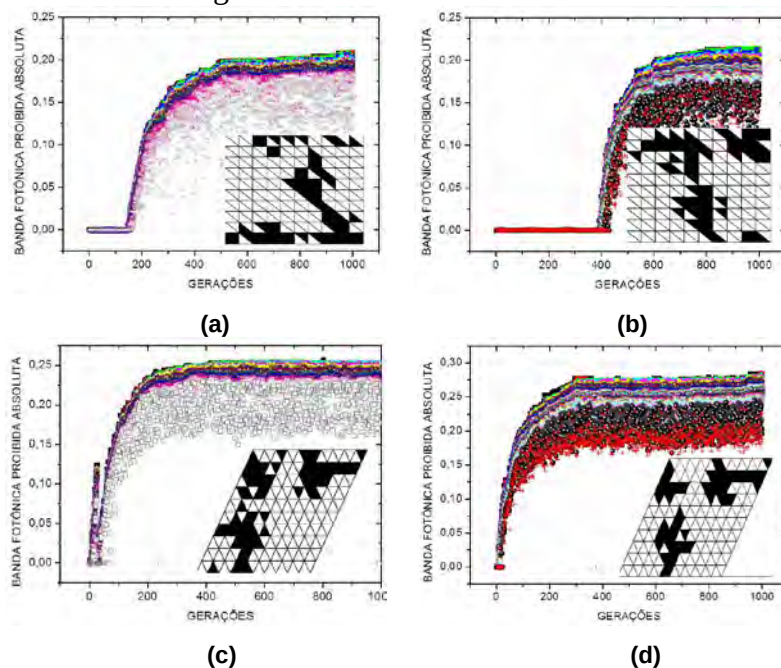


Figura 4. Evolução do melhor indivíduo do cristal: com arranjo quadrado (a) população com 20 indivíduos e (b) população com 50 indivíduos e com arranjo triangular: (c) população com 20 indivíduos e (d) população com 50 indivíduos.

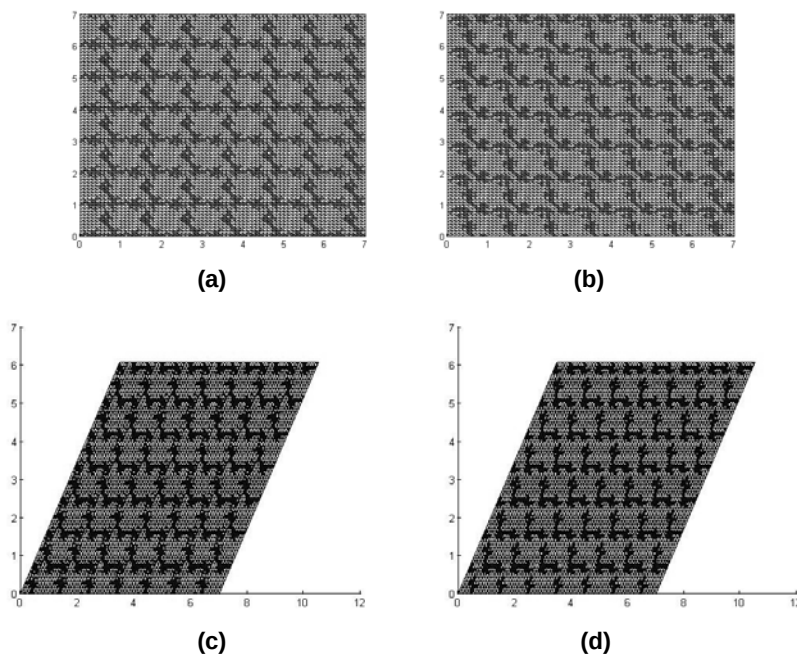


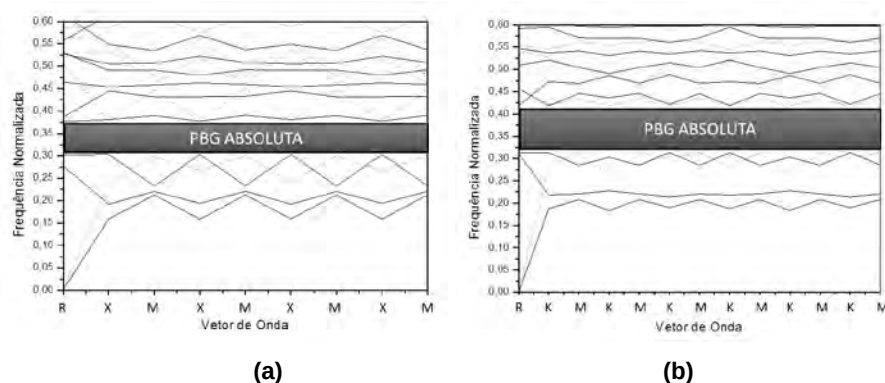
Figura 5. Cristais Fotônicos: arranjo quadrado (a) 20 indivíduos (b) 50 indivíduos e arranjo triangular (c) 20 indivíduos (d) 50 indivíduos.

A mutação utilizada permite explorar um espaço de busca maior. Além disso, abrange muitas possibilidades de configurações através do tamanho da população utilizada. As geometrias resultantes, indicadas nas Figuras 5 e 7, são assimétricas. O mesmo comportamento pode ser visto em [Malheiros, Esquerre e Figueroa 2011].

Tabela 1. PBG Absoluta para o Arranjo Quadrado e Arranjo Triangular

Cristal com Arranjo Quadrado			
Método utilizado em [Malheiros 2011]	20 indivíduos	SIA	50 indivíduos
18,01%	21,24%	21,27%	
Cristal com Arranjo Triangular			
Método utilizado em [Malheiros 2011]	20 indivíduos	SIA	50 indivíduos
23,49%	25,67%	28,45%	

Diferentemente de círculos e elipses, que são as mais tradicionais analisadas, as estruturas resultantes não são triviais, tornando o processo de fabricação complexo. No entanto, se PBGs em modos de ordens superiores forem utilizadas, a constante de arranjo seria maior e a robustez e tolerância de fabricação melhorada [Malheiros, Esquerre e Figueroa 2011]. Neste trabalho as PBGs são de modos de ordens inferiores para serem comparados com os resultados encontrados na literatura [Malheiros, Esquerre e Figueroa 2011]. O mesmo refinamento é considerado para o arranjo quadrado e triangular.

**Figura 6. Dispersão (a) arranjo quadrado (b) arranjo triangular**

5. Conclusões

A otimização de bandas fotônicas proibidas absolutas de cristais fotônicos compostos de telúrio e ar, utilizando o método dos elementos finitos em conjunto com o algoritmo de sistema imunológico artificial é apresentado. Os arranjos cristalinos utilizados foram quadrados e triangulares e as geometrias resultantes são consideradas como sendo buracos de ar no telúrio. Existem limitações na fabricação das geometrias encontradas por serem mais complexas se comparado com hastes dielétricas de telúrio no ar ou buracos de ar num substrato de telúrio [Malheiros, Esquerre e Figueroa 2011]. Mas, estas limitações podem ser superadas se as PBGs são projetadas para operar em frequências menores onde a constante de arranjo é aumentada.

Os autores gostariam de agradecer o suporte financeiro da UFBA, IFBA, INCT Fotonicom, FAPESB, CAPES e CNPq Processo: 302390/2009-0.

Referências

Joannopoulos, J. D., Johnson, S. G., Winn, J. N. e Meade, R. D. (2008) "Photonic Crystals: Molding the Flow of Light", 2ª ed., Princeton University Press.

- Malheiros, G. N. S., Esquerre, V. F. R. e Figueroa, H. E. H. (2011) “A Strategy of Search e Refinement by GA in 2D Photonic Crystals with Absolute PBG”, Id: JQE-132416-2010, n. de pgs: 08. IEEE Journal of Quantum Electronics.
- Li, Z.-Y., Gu, B.-Y. e Yang, G.-Z. (1998) “Large absolute band gap in 2- D anisotropic photonic crystals”, *Phys. Rev. Lett.*, vol. 81, no. 12, pp. 2574–2577, September.
- Li, Z.-Y., Gu, B.-Y. e Yang, G.-Z. (1999) “Improvement of absolute band gaps in 2-D photonic crystals by anisotropy in dielectricity”, *Eur. Phys. J. B*, vol. 11, no. 1, pp. 65–73, September.
- Zaccaria, R. P., Verma, P., Kawaguchi, S., Shoji, S. e Kawata, S. (2008) “Manipulating full photonic band gaps in 2-D birefringent photonic crystals”, *Opt. Exp.*, vol. 16, no. 19, pp. 14812–14820, September.
- Rezaei, B., Khalkhali, T. F., Vala, A. S. e Kalafi, M. (2009) “Absolute band gap properties in 2-D photonic crystals composed of air rings in anisotropic tellurium background”, *Opt. Commun.*, vol. 282, no. 14, pp. 2861–2869, July.
- Kurt, H. e Citrin, D. S. (2005) “Annular photonic crystals”, *Opt. Exp.*, vol. 13, no. 25, pp. 10316–10326, December.
- Rodriguez-Esquerre, V. F., Hernández-Figueroa, H. E., e Koshiba, M. (2006) “Modeling of complex structures devices and applications”, in *Telecommunications: Advances and Trends in Transmission, Networking and Applications*, Cavalcante, et al, Eds. Fortaleza, Brazil: Univ. Fortaleza Press, c. 8, pp. 173–187.
- Lin, M.-C. e Jao, R.-F. (2007) “Finite element analysis of photon density of states for 2- D photonic crystals with in-plane light propagation”, *Opt. Exp.*, vol. 15, no. 1, pp. 207–218, January.
- Castro, L. N. (2006) “Fundamentals of Natural Computing Basic Concepts, Algorithms and Applications”, Chapman & Hall/CRC, London, U.K.
- Silva-Santos, C. H.; Rodríguez-Esquerre, V. F. e Hernández-Figueroa, H. E. (2010a) “An Artificial Immune System for Optical Fiber Based Directional Couplers Multiplexer/Demultiplexers Design”, In: *LAOP 2010 - Latin America Optics and Photonics Conference*, 2010, Recife, PE.
- Silva-Santos, C. H., Gonçalves, M. S. e Hernández-Figueroa, H. E. (2010b) “Designing Novel Photonic Devices by Bio-Inspired Computing”, *IEEE Photonics Technology Letters*, v. 22, pp. 1777-1779.
- Quiñónez, F., Menezes, J. W., Cescato, L., Rodríguez-Esquerre, V. F., Hernández-Figueroa, H. e Mansano, R. D. (2006) “Band gap of hexagonal 2-D photonic crystals with elliptical holes recorded by interference lithography”, *Opt. Exp.*, vol. 14, no. 11, pp. 4873–4879, May.
- Shi, Y. (2010) “A compact polarization beam splitter based on a multimode photonic crystal waveguide with an internal photonic crystal section”, *Progress Electromagn. Res.*, vol. 103, pp. 393–401.
- Dourado-Sisnando, A., Santos, F. N., Rodríguez-Esquerre, V. F. e Malheiros, G. N. S. (2012) “Comparación de Algoritmos Evolutivos Aplicados en La Optimización de Bandas Fotónicas Prohibidas en Cristales Fotónicos Bidimensionales”, In: *COIN-TEC 2012 – IV Congreso Internacional de Tecnología 2012*, Lima, Peru.

SEDOR - Sistema Especialista para Diagnóstico Odontológico baseado em Radiografias

Renato Borges¹, Gabriela Saddi², Rogerio Salvini¹, Fernanda Yamamoto³

¹Instituto de Informática – Universidade Federal de Goiás (INF-UFG) – Goiânia – GO

²Associação Brasileira de Odontologia – Seção Goiás (ABO) – Goiânia – GO

³Faculdade de Odontologia – Universidade Federal de Goiás (FO-UFG) – Goiânia – GO

{renatojunior, rogeriosalvini}@inf.ufg.br

Abstract. *Obtaining accurate diagnosis in Dentistry has always been a challenge for professionals, especially those who have just started their careers. The knowledge required to get to a correct diagnosis, coupled with the variety of lesions with similar clinical and radiographic findings, led to the creation of systems for clinical decision support, called Expert Systems. In this work we presented an expert system to aid the process of diagnosing benign radiolucent lesions, that integrates a database based on radiographic features and a user-friendly web interface.*

Resumo. *A obtenção de diagnóstico preciso na Odontologia sempre foi um desafio para os profissionais da área, especialmente aqueles em início de carreira. O conhecimento necessário para se chegar a um diagnóstico correto, associada à variedade de lesões com aspectos clínicos e radiográficos semelhantes, levou à necessidade da criação de sistemas de apoio à decisão clínica, chamados de Sistemas Especialistas. Neste trabalho apresentamos um sistema especialista de apoio à decisão clínica para auxiliar o diagnóstico de lesões ósseas benignas radiolúcidas, que integra uma base de dados baseada em características radiográficas e uma interface web de fácil utilização.*

1. Introdução

O principal aspecto a ser analisado no atendimento do paciente é a necessidade de se garantir um diagnóstico preciso. O planejamento do tratamento e o prognóstico são fundamentados no diagnóstico e, por esta razão, faz-se necessário uma correta diferenciação das diversas doenças que afetam o complexo oral e maxilofacial. Devido à similaridade na apresentação clínica de várias entidades patológicas, embora sejam distintas suas patogêneses e etiologias, a decisão do diagnóstico torna-se uma tarefa árdua [Neville *et al.*, 2009].

O exame radiográfico é uma ferramenta imprescindível na Odontologia e constitui um recurso auxiliar do diagnóstico, pois apesar do exame clínico ser imperativo, as radiografias podem contribuir para detectar doenças ósseas que não poderiam ser visualizadas pelo exame físico.

O uso de sistemas computacionais pelos profissionais de saúde tem crescido significativamente nos últimos anos. Os avanços tecnológicos, principalmente da informática, têm proporcionado considerável auxílio na área do diagnóstico [Rudin, 1994]. Para profissionais isolados ou menos experientes, esses sistemas de auxílio

diagnóstico são benéficos por fornecerem uma segunda opinião [Ralls et al., 1986]. Isto é especialmente verdadeiro nas situações em que pequenas diferenças no julgamento de diagnóstico podem resultar em tratamentos radicalmente diferentes.

Neste trabalho apresentamos um sistema especialista para auxiliar o processo de diagnóstico de lesões radiolúcidas benignas baseado em características radiográficas. Este sistema, denominado SEDOR (Sistema Especialista para Diagnóstico Odontológico baseado em Radiografia), foi construído a partir de estudos científicos publicados sobre várias lesões, extraíndo-se as características de cada uma delas. A partir desta base científica as lesões foram colocadas de forma analítica em uma base de conhecimento utilizada pelo sistema para inferir hipóteses diagnósticas para auxiliar o cirurgião-dentista.

Este artigo está organizado da seguinte forma. Na seção seguinte é apresentada uma fundamentação teórica básica sobre sistemas especialistas, e alguns trabalhos importantes sobre sistemas de diagnóstico baseados em radiografias são brevemente comentados. Na Seção 3 detalhamos o desenvolvimento do SEDOR. A avaliação e validação do sistema são mostradas na Seção 4. Por fim, na Seção 5, são apresentadas as considerações finais do trabalho e as perspectivas para trabalhos futuros.

2. Fundamentação Teórica e Trabalhos Relacionados

2.1. Sistemas Especialistas

Não existe uma definição exata do que é um Sistema Especialista (SE). De acordo com [Levine *et al.*, 1988] “um sistema de inteligência artificial criado para resolver problemas em um domínio particular é chamado de Sistema Especialista”. Em [Bielawski e Lewand, 1988] é proposto uma definição mais formal: “um Sistema Especialista é um programa que simula o desempenho de um especialista humano em um domínio ou campo específico”. Pode-se dizer que SE são sistemas que utilizam de conhecimento empírico baseado na experiência de um profissional humano em uma área de conhecimento para resolver problemas deste domínio, procurando chegar a uma conclusão utilizando de forma semelhante a mesma forma de raciocínio que o especialista humano. Dentre os sistemas especialistas mais conhecidos, podemos citar: MYCIN, PROSPECTOR e DENTRAL [Alty, 1985].

Podemos dividir a arquitetura de sistemas especialistas em três componentes principais: base de conhecimento, motor de inferência e interface de usuário. A base de conhecimento é o componente responsável por armazenar as informações a respeito do domínio do problema que se deseja resolver. Estas informações são compostas por fatos e regras, sendo possível obtê-las por meio do conhecimento de especialistas humanos ou de bases de dados que muitas vezes são utilizadas por estes profissionais. O motor de inferência consulta a base de conhecimento para determinar uma ação a ser realizada. Ele é o mecanismo responsável por manipular o conhecimento já presente na base de conhecimento, podendo inferir novos fatos ou chegar a conclusões a respeito do problema. A interface de usuário é uma representação gráfica que permite o usuário alvo interagir com o sistema. Nela são exibidas as respostas, as previsões e as sugestões obtidas pelo motor de inferência.

2.2. Diagnóstico Odontológico baseado em Radiografias

O primeiro exemplo de sistema de computador para uso em diagnóstico oral foi desenvolvido por [Leonard *et al.*, 1973] que baseava-se em um modelo de

reconhecimento de padrões linear. Ele buscava fornecer o diagnóstico e um plano de tratamento para dor e disfunção miofacial - um tipo de dor craniofacial. Firriolo criou um protótipo de sistema para ajudar na análise e diagnóstico baseados em radiografias panorâmicas denominado CAREOP (*Computer Assisted Radiographic Evaluation of Oral Pathology*) [Firriolo, 1986 *apud* Siegel *et al.*, 1993].

Um dos principais sistemas de apoio à decisão diagnóstica baseadas em radiografias foi desenvolvido por [White, 1989] denominado ORAD (*Oral Radiographic Differential Diagnosis*), que utiliza métodos estatísticos para avaliar as características clínicas e radiográficas de pacientes com lesões intraósseas com intuito de auxiliar na sua identificação. Foram descritas, a princípio, 98 lesões dos maxilares pelas suas prevalências e distribuição por idade, sexo, raça, presença de dor, número, tamanho e localização das lesões, associação com dentes, expansão, loculação, bordas, conteúdo e efeitos nos dentes adjacentes.

Hubar *et al.* desenvolveram o programa COMRADD (*Computerized Radiographic Differential Diagnosis*), que inclui tanto alterações ósseas quanto anormalidades do elemento dentário em radiografias. No total, 125 patologias orais foram categorizadas dentro de uma de cinco diferentes classes, sendo as maiores: alterações radiográficas do elemento dentário, lesões radiográficas de tecidos moles e lesões radiolúcidas, mistas e radiopacas do crânio, maxila e mandíbula [Hubar *et al.*, 1990].

Wiener *et al.* descrevem um sistema para ajudar no diagnóstico de quatro cistos e oito tumores de origem odontogênica, fundamentado em métodos estatísticos baseado em achados clínicos e radiográficos do paciente [Wiener *et al.*, 1986].

3. SEDOR

3.1. Framework

O desenvolvimento do SEDOR se deu utilizando a linguagem de programação Java em conjunto com um *framework* de SE. A escolha quanto ao *framework* utilizado se motivou principalmente por dois fatores: possibilidade de integração com outras linguagens e que possuísse uma documentação vasta e de qualidade.

O CLIPS (*C Language Integrated Production System*) é um *framework* de SE projetado para facilitar o desenvolvimento de softwares capazes de modelar o conhecimento humano ou uma perícia. Originalmente desenvolvido pela *Software Technology Branch* (STB), NASA/Lyndon B. Johnson Space Center, hoje já é usado por milhares de pessoas no mundo todo [Giarratano, 2007].

Outro *framework* é o JESS [Friedman-Hill, 2008], uma biblioteca que integra o CLIPS e a linguagem Java, por meio de classes e métodos em Java que fazem chamadas ao motor do CLIPS. Além de estar diretamente ligado ao Java, o JESS possui algumas adições em relação ao CLIPS. Por exemplo, ele permite adicionar objetos instanciados em Java a sua memória de trabalho. Além disto, também é possível criar e manipular objetos Java e transferir valores entre o código JESS e Java. O JESS também possui uma ampla documentação e uma licença de uso acadêmico.

Desta forma, o JESS foi a escolha final como o *framework* para desenvolvimento do SEDOR. A sua integração com a linguagem Java nos permitiu separar a construção da base de conhecimento e o motor de inferência, que processará o diagnóstico, de outros módulos do sistema, como a interface gráfica, por exemplo.

3.2. Representação do Conhecimento

Para formar a base de conhecimento do sistema tivemos a participação ativa de uma professora da Faculdade de Odontologia da Universidade Federal de Goiás (FO-UFG) especialista em Patologia Bucal, Diagnóstico, Radiologia Odontológica e Imaginologia.

Utilizamos planilhas eletrônicas para expressar o conhecimento do especialista. Optamos por esta ferramenta pela facilidade de uso e interação, e também por ajudar na conversão nas regras de produção. Uma planilha modelo foi usada para gerar as planilhas de todas as lesões estudadas. Seu principal objetivo é criar um *template* do JESS que permita preencher as diversas propriedades para um diagnóstico.

Além do gênero e faixa etária do paciente, o especialista selecionou os seguintes atributos radiográficos para se fazer o diagnóstico de uma lesão: Densidade, Localização, Limites, Forma, Efeitos nos tecidos adjacentes e Interior da lesão. Para cada atributo, o especialista também indicou um conjunto de valores possíveis que o atributo poderia possuir.

Na Figura 1 é mostrada uma parte da planilha modelo. A coluna A representa o nome do atributo. A coluna B é o peso que aquele atributo exerce no diagnóstico final. A coluna C é o percentual relativo ao peso do atributo. A coluna D representa os valores de cada atributo. As colunas E e F são semelhantes as colunas B e C, porém referem-se aos pesos e probabilidades dos valores, respectivamente. Por fim a coluna G refere-se ao tipo de cada atributo que podem ser: “u” (univalorados, apenas um valor pode ser selecionado pelo usuário) ou “m” (multivalorados, mais de um valor pode ser selecionado).

	A	B	C	D	E	F	G
1	NOVA_LESAO						
2							
3	Atributo	Peso	%	Valores	Peso	%	Tipo
4	GENERO			FEMININO			u
5	GENERO			MASCULINO			
6	FAIXA_ETARIA			0-9			u
7	FAIXA_ETARIA			10-19			
8	FAIXA_ETARIA			20-29			
14	DENSIDADE			RADIOLUCIDA			u
15	DENSIDADE			MISTA			
16	DENSIDADE			RADIOPACA			
17	LOCALIZACAO			MAXILA_ANTERIOR			m
18	LOCALIZACAO			MAXILA_POSTERIOR			
19	LOCALIZACAO			MANDIBULA_ANTERIOR			
38	LIMITES			DEFINIDOS			u
39	LIMITES			DEFINIDOS_(CORTICALIZADO)			
40	LIMITES			DISCRETAMENTE_DEFINIDOS_(CORTICAL_POUCO_NITIDA)			
43	FORMA			UNILOCLAR			m
44	FORMA			MULTILOCLAR			
45	FORMA			OVALADA_OU_ARREDONDADA			
48	EFEITOS_NOS_TECIDOS_ADJACENTES			EXPANSAO_DAS_CORTICAIS_OSSEAS_(ABAUAMENTO)			m
49	EFEITOS_NOS_TECIDOS_ADJACENTES			ADELGACAMENTO_DAS_CORTICAIS_OSSEAS			
50	EFEITOS_NOS_TECIDOS_ADJACENTES			REABSORCAO_RADICULAR_DO(S)_DENTE(S)_ENVOLVIDO(S)			
81	INTERIOR_DA_LESAO			RADIOLUCIDA			u
82	INTERIOR_DA_LESAO			RADIOLUCIDA_COM_FOCOS_DE_CALCIFICACAO			
83	INTERIOR_DA_LESAO			RADIOLUCIDA_COM_FOCOS_DE_CALCIFICACAO_PROXIMO_A_COROA_DENTARIA			
84	INTERIOR_DA_LESAO			MISTA			
85	INTERIOR_DA_LESAO			RADIOPACA			
86	INTERIOR_DA_LESAO			SEPTACOES			

Figura 1. Planilha modelo

As demais planilhas foram baseadas na planilha modelo e expressam as características de cada lesão. Foram registradas 29 lesões, gerando 29 planilhas. Para cada lesão, o especialista atribuiu um peso para cada atributo referente à sua importância no diagnóstico da lesão. Estes pesos foram atribuídos pelo especialista baseados em pesquisas na literatura específica e na experiência de casos clínicos. Da mesma forma foram atribuídos pesos para os valores dos atributos. Se um valor de atributos ficasse com valor zero isso indicaria que aquele valor não ocorre em um atributo para uma

determinada lesão. A partir dos pesos dados, os percentuais relativos, tanto dos atributos quanto dos valores do atributo, são calculados automaticamente e serão usados na avaliação de um diagnóstico pelo sistema.

Outra informação contida nas planilhas é a frequência de cada lesão, expresso como um valor real entre 0 e 1, na linha 1 e coluna B da Figura 1. Ela foi utilizada para expressar a raridade das lesões, pois algumas ocorrem com mais frequência que outras.

3.3. Base de Conhecimento e Motor de Inferência

O JESS encontra-se na base do sistema. Ele faz a construção das regras e fatos iniciais na base de conhecimento e é o seu motor de inferências quem faz o processamento para determinar o melhor diagnóstico de uma entrada de sintomas (fatos) presentes na radiografia de um paciente. O código Java lê as planilhas e cria as regras de produção no formato que é compreendido pelo JESS.

O método utilizado pelo motor de inferência para dar o diagnóstico de uma lesão é conhecido como “soma de pesos”. Uma função linear determina a probabilidade de cada lesão individual por meio da soma dos pesos relacionados a cada característica ou atributo que contribui com uma parcela da probabilidade final da lesão. Após receber a entrada do usuário com as características presentes em uma radiografia, o sistema calcula para cada lesão a probabilidade resultante, somando os pesos de cada valor escolhido pelo usuário em cada uma delas. Por fim, as probabilidades são ordenadas para determinar as lesões mais prováveis.

Foi gerado uma regra para cada valor possível em uma lesão, de forma que, o antecedente da regra irá validar se o valor foi escolhido pelo usuário, e então, a regra é ativada e o consequente da regra adiciona a probabilidade correspondente ao peso deste valor na probabilidade final da lesão.

3.4. Interface Gráfica

Utilizando JSP (*Java Server Pages*) e *Servlets*, presentes na linguagem Java e utilizados para criação de sistemas Web, foi criada uma interface gráfica para a Web que pudesse executar em um navegador. Essa escolha se deu devido à grande difusão dos sistemas Web, eliminando problemas com instalação ao mesmo tempo que facilita uso do sistema.

O sistema possui três diferentes telas, acessadas por meio de um menu com as opções: Diagnóstico, Lesões e Testes. A tela de Diagnóstico é apresentada na Figura 2, é nela que o usuário seleciona as entradas para o sistema realizar o diagnóstico.

A Figura 3 exibe o resultado do diagnóstico após submeter os valores da tela anterior. As 5 mais prováveis lesões são exibidas e logo abaixo também é mostrado os valores que foram escolhidos de cada atributo.

No menu Lesões, é possível adicionar uma nova lesão por meio de uma planilha no formato de arquivo *.csv*, utilizado para representar uma lesão, além disto ela também lista todas as lesões já presentes no sistema. A tela de Testes permite realizar diagnósticos a partir de planilhas que descrevem as características de uma radiografia. Esta opção é semelhante a de Diagnóstico, porém ao utilizar planilhas prontas, ela permite realizar o diagnóstico de várias entradas de uma única vez, armazenando e exibindo o resultado final.




SEDOR

Sistema Especialista para Diagnóstico Odontológico baseado em Radiografia

Diagnóstico

Lesões

Testes

Forma

Select options

Efeitos nos tecidos adjacentes

2 selected

Densidade

Radiolucida

Faixa etaria

0-9

Interior da lesao

1 selected

Localizacao

Figura 2. Tela de diagnóstico




SEDOR

Sistema Especialista para Diagnóstico Odontológico baseado em Radiografia

Diagnóstico

Lesões

Testes

Diagnóstico

1 - Cisto radicular: 61,2%
2 - Granuloma dentario: 60,3%
3 - Abscesso periapical: 41,5%
4 - Cisto periodontal lateral: 27,0%
5 - Cisto botriode: 13,8%

Sintomas

Forma

- Ovalada ou arredondada

Efeitos nos tecidos adjacentes

- Perda da lamina dura

Densidade

- Radiolucida

Faixa etaria

- 10-19

Interior da lesao

- Radiolucida

Figura 3. Tela de resultados

4. Testes e Validação

Foram selecionadas 10 radiografias panorâmicas obtidas por um aparelho radiográfico convencional. As radiografias selecionadas tinham o diagnóstico definitivo registrado em seu prontuário, cujo exame anatomopatológico foi realizado pela disciplina de Patologia Oral do Departamento de Ciências Estomatológicas da FO-UFG. Todas as radiografias apresentavam uma lesão benigna de densidade radiolúcida e tinham como diagnóstico

final umas das 29 lesões ósseas contidas na base de conhecimento do sistema.

Foram convidados 3 observadores, sendo um acadêmico do 6º período do curso de Odontologia, um Radiologista e um Patologista, os dois últimos com mais de cinco anos de experiência clínica. Cada observador analisou as 10 radiografias selecionadas e tinha a opção de escrever até três hipóteses diagnósticas. Depois, os dados analisados pelo observador foram inseridos no SEDOR e obtidos uma lista de cinco prováveis diagnósticos pelo programa, os quais foram considerados corretos se a lesão aparecesse até o quinto lugar.

Para avaliar as facilidades do sistema, foi também realizado um questionário com seis perguntas referentes ao sistema, as quais foram respondidas por cada um dos observadores.

As avaliações visando a validação do programa foram realizadas pelos três observadores. Os valores percentuais de acertos dos observadores convidados e do sistema são mostrados na Tabela 1. As opiniões sobre a usabilidade do sistema são apresentadas na Tabela 2.

Tabela 1. Valores percentuais de acertos dos observadores e do SEDOR

	Acerto sem o SEDOR	Acerto do SEDOR
Observador 1	50%	40%
Observador 2	60%	60%
Observador 3	60%	70%

Tabela 2. Número de respostas (SIM ou NÃO) que o sistema obteve dos observadores para cada pergunta sobre sua usabilidade

Perguntas	SIM	NÃO
Você acha que o programa é de fácil manuseio?	3	0
Você conseguiu utilizar o programa sem a necessidade de explicações adicionais?	2	1
O programa apresenta aspectos radiográficos não avaliados inicialmente?	2	1
Você utilizaria o programa na prática clínica?	3	0
Você adquiriria o programa?	3	0
O programa possui uma boa interface?	2	1

5. Conclusão

Os sistemas de apoio à decisão clínica têm mostrado bons resultados quando utilizados para incrementar o conhecimento dos profissionais da Odontologia quando da realização de diagnósticos. É evidente que tais sistemas não são capazes de sozinhos chegarem a um diagnóstico preciso, mas se converteram em um importante instrumento de apoio, principalmente para profissionais sem grande experiência em diagnósticos.

Consideramos que o SEDOR atingiu o propósito para o qual foi concebido. Os percentuais de acerto entre o SEDOR e os observadores são comparáveis em relação ao diagnóstico registrado no prontuário das radiografias selecionadas. Além disso, houve

um alto grau de aceitação do sistema pelos observadores.

No prosseguimento deste trabalho, iremos refinar os pesos das planilhas das lesões já existentes, principalmente aquelas onde há maior dificuldade no acerto do diagnóstico. Além disso, pretendemos aumentar o número de lesões da base de conhecimento para abranger outros grupos de lesões e aplicar novas estratégias de representação. Novos testes serão executados com um número maior de casos e de observadores especialistas, permitindo-nos fazer mais tipos de comparações.

Referências

- Alty, J. "Use of expert systems", *Computer-Aided Engineering Journal*, 2(1):2-9, February, 1985.
- Bielawski, L.; Lewand, R. "Expert Systems Development", QED Information Sciences, Inc., 1988.
- Friedman-Hill, E. "Jess the Rule Engine for the Java Platform", Sandia National Laboratories, November, 2008.
- Giarratano, J. "CLIPS User's Guide", University of Houston-Clear Lake, december 2007.
- Hubar, J.S.; Manson-Hing, L.R.; Heaven, T. "COMRADD: Computerized Radiographic Differential Diagnosis", *Oral Surgery, Oral Medicine, Oral Pathology, Alabama*, v. 69, n. 2, p.263-265, 1990.
- Leonard, M.S. *et al.* "Automated Diagnosis and Treatment Planning of Craniofacial Pain", *J. Dent. Res., Florida*, v. 53, n. 5, p.1155-1159, 1974.
- Levine, R.L.; Drang, D. E.; Edelson, B. "A Comprehensive Guide to AI and EXPERT SYTEMS: Turbo Pascal Edition", McGraw-Hill, 1988.
- Neville, B.W. *et al.* "Patologia Oral e Maxilofacial", 3 ed. Rio de Janeiro, Elsevier, 2009.
- Ralls, S.A.; Cohen, M.E.; Southard, T.E. "Computer-Assisted Dental Diagnosis", *Dental Clinic of North America*, v. 30, n. 4, p.695-712, 1986.
- Rudin, Jonathan L. "DART (Diagnostic Aid and Resource Tool): A Computerized Clinical Decision Support System for Oral Pathology", *Compendium*, v. 15, n. 11, p.1316-1328, 1994.
- Siegel, M.A.; Firriolo, F.J.; Finkelstein, M.W. "Computer Applications in Oral Diagnosis", *Dental Clinics of North America*, v. 37, n. 1, p.113-131, 1993.
- White, S.C. "Computer-aided differential diagnosis of oral radiographic lesions", *Dentomaxillofac. Radiol.*, v. 18, p.53-59, 1989.
- Wiener, F.; Laufer, D.; Ribak, A. "Computer-aided diagnosis of odontogenic lesions", *International Journal of Orall and Maxillofacial Surgery, Israel*, v. 15, p.592-596, 1986.

Avaliação do Desempenho de Esquemas de Cruzamento em Algoritmos Genéticos na Otimização de Bandas Proibidas em Cristais Fotônicos

Francismari Noronha dos Santos^{1,2}, Anderson Dourado Sisnando²,
Vitaly Felix Rodriguez Esquerre², Gilliard Nardel Malheiros Silveira³

¹Departamento de Eletrotécnica, Instituto Federal de Educação, Ciência e Tecnologia da Bahia (IFBA) - Salvador, BA - Brasil.

²Programa de Pós-graduação em Engenharia Elétrica, Universidade Federal da Bahia (UFBA) - Salvador, BA - Brasil.

³Departamento de Microondas e Óptica, Universidade Estadual de Campinas (UNICAMP) - Campinas, SP - Brasil.

mari.noronhas@gmail.com, anderson.sisnando@ufba.br,
vitaly.esquerre@ufba.br, gilliardnardel@yahoo.com

Abstract. *The aim of this work is the assessment of two crossover schemes in genetic algorithms applied for the optimization of the band gap of photonic crystal structures with two dimensional periodicity and considering the square and triangular lattices. It was considered the band gap between the first and the second transverse magnetic modes, denoted TM_{12} . The propagation properties of the crystals are obtained by an efficient finite element method in the frequency domain.*

Resumo. *O objetivo deste trabalho é a avaliação de dois esquemas de cruzamento em algoritmos genéticos aplicados na otimização da banda proibida em estruturas de cristal fotônico com periodicidade em duas dimensões e considerando os arranjos quadrado e triangular. Considerou-se a banda proibida entre o primeiro e o segundo modos transversais magnéticos, denotado TM_{12} . As propriedades de propagação dos cristais são obtidas pelo método de elemento finito eficiente no domínio da frequência.*

1. Introdução

Dispositivos baseados em cristais fotônicos são de grande aplicação nos sistemas de comunicações ópticas, tais como: demultiplexadores, cavidades ressonantes, filtros entre outros [Joannopoulos et al. 2008]. Elas apresentam bandas proibidas que são faixas de frequências, nas quais a luz não se propaga em nenhuma direção no cristal. Esta propriedade é explorada no projeto de novos dispositivos. Uma forma de otimizar esses cristais implica em obter estruturas que apresentem uma maior banda fotônica proibida para os modos de propagação desejados, o que permitiria uma maior largura de banda de operação.

Neste trabalho, a banda fotônica proibida de cristais fotônicos bidimensionais (arranjos quadrados e triangulares) é otimizada, maximizando-a para os primeiros modos transversais magnéticos, denominada TM_{12} através de algoritmos genéticos utilizando o cruzamento uniforme, com o objetivo de comparar com os resultados obtidos em [Malheiros e Esquerre 2007], no qual o cruzamento de dois pontos foi utilizado. A

proposta de utilizar o cruzamento uniforme ao invés do de dois pontos é achar novos ótimos, já que esse tipo de cruzamento poderá aumentar o espaço de busca, por possibilitar uma maior diversidade na população.

O algoritmo genético, assim como outros algoritmos evolutivos, tem se mostrado uma boa alternativa na solução de problemas eletromagnéticos, que é o caso do problema apresentado neste trabalho. [Malheiros e Esquerre 2007], [Haupt e Werner 2007] e [Preble e Lipson 2005].

São realizadas simulações considerando uma população de 20 e 50 indivíduos (cromossomos) em cada geração, a evolução acontece durante 1000 gerações para cada tipo de cruzamento (dois pontos e uniforme).

2. Cristais Fotônicos Bidimensionais

Em um cristal fotônico, a luz que atravessa o mesmo é espalhada pela geometria periódica sendo refratada e parcialmente refletida nas interfaces entre os materiais. O padrão resultante das ondas será reforçado ou cancelado dependendo do comprimento de onda da luz, o índice de refração dos materiais, sua direção de propagação através do cristal, e o tamanho e arranjo das estruturas (geometria), e a partir dessas iterações, aparece a banda fotônica proibida.

Por apresentarem simetrias, os cristais podem ter a análise dos campos e modos eletromagnéticos a partir de células unitárias com condições de contorno periódicas. A partir desta célula é obtida a primeira zona de Brillouin, a qual define a região do espaço dos vetores de onda que determinam os modos de propagação (Ver Figura 1).

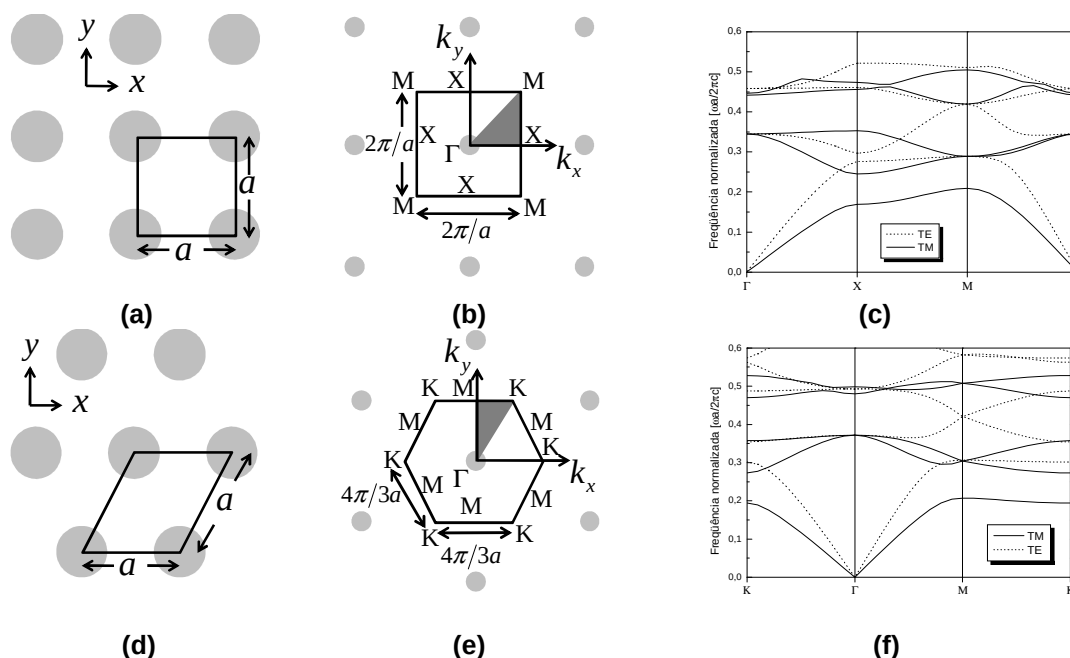


Figura 1. (a) Célula unitária do arranjo quadrado [Malheiros, Esquerre e Figueroa 2011]. (b) região de Brillouin [Malheiros, Esquerre e Figueroa 2011]. (c) estrutura de banda fotônica para um arranjo quadrado formado por cilindros dielétricos com $n=3,6$ em ar e $r/a=0,35$. (d) célula unitária do arranjo triangular [Malheiros, Esquerre e Figueroa 2011]. (e) região de Brillouin [Malheiros, Esquerre e Figueroa 2011]. (f) estrutura de banda fotônica para um arranjo triangular formado por cilindros dielétricos com $n=3,6$ em ar e $r/a=0,33$.

A região sombreada (triângulo) nas Figuras 1.b e 1.e correspondem à região irredutível de Brillouin limitadas por $\Gamma=(0,0)\pi/a$, $X=(1,0)\pi/a$, $M=(1,1)\pi/a$ para o arranjo quadrado e por $\Gamma=(0,0)\pi/a$, $K=(2/3, 2\sqrt{3}/3)\pi/a$, e $M=(0, 2\sqrt{3}/3)\pi/a$ para o arranjo triangular.

Nas Figuras 1.c e 1.f são apresentados o comportamento dos vetores de onda para os modos TE e TM num cristal com arranjo quadrado e triangular, fica visível que para estes casos existe uma Banda Fotônica Proibida entre as frequências normalizadas de 0,2-0,24 e 0,2-0,27, respectivamente. As geometrias são geradas no programa computacional GID¹ e as bandas proibidas são calculadas através do método dos elementos finitos no domínio da frequência [Malheiros, Esquerre e Figueroa 2011]. Nesse método, o sistema matricial resultante é um problema de autovalores e autovetores. Neste trabalho foi considerado um cristal composto de ar e silício, cujos índices de refração são $n=1,0$ e $n=3,476$, respectivamente. As malhas correspondentes às células unitárias são apresentadas na Figura 2. Cada triângulo representa um gene do cromossomo e ele poderá ser de ar ou silício.

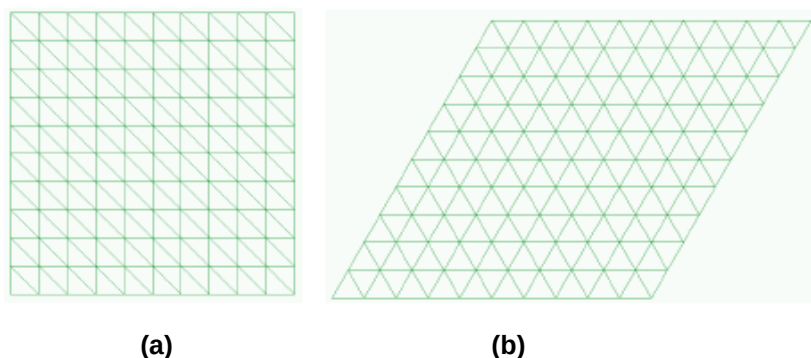


Figura 2. Célula unitária dos cristais fotônicos dividida em 200 triângulos correspondente ao (a) arranjo quadrado e (b) arranjo triangular, respectivamente [Malheiros e Esquerre 2007].

3. Algoritmos Genéticos

Para o problema de otimização da banda fotônica proibida o algoritmo genético se mostrou uma boa opção, pois o cálculo das bandas proibidas requer de tempo de computação relativamente longo, o que torna inviável realizar uma varredura em todo o universo de possibilidades, que corresponde a 2^{200} .

O algoritmo genético utilizado neste trabalho é constituído das seguintes etapas (ver Figura 3):

a) Inicialização: nesta etapa é gerada a população inicial. Considerou-se uma célula unitária para cada malha, como mostrado na Figura 2. Elas foram divididas em 200 elementos triangulares, resultando em 441 nós, cada triângulo representa um gene do cromossomo. O indivíduo é representado como sequência binária, se o triângulo possuir valor zero ele é composto por ar e caso seja um é composto por silício. Portanto, cada cromossomo é composto de 200 genes (variáveis). As simulações foram feitas com 20 e 50 indivíduos.

¹ O GID é um programa computacional que utiliza o método numérico na análise de dispositivos. Informações sobre o programa podem ser obtidas em: <http://www.gidhome.com>.

b) Avaliação: consiste em avaliar as aptidões das soluções, a função que avalia os indivíduos (cromossomos), é realizada através da seguinte expressão:

$$\text{Função Objetivo} = \frac{\text{Frequência}_{\text{superior}} - \text{Frequência}_{\text{inferior}}}{\text{Frequência}_{\text{média}}} \cdot 100\% \quad (1)$$

c) Seleção: nesta etapa os indivíduos são selecionados para reprodução, a probabilidade de uma determinada solução ser selecionada é proporcional à sua aptidão, foi utilizada a seleção por *Roulette Wheel* [Affenzeler et al. 2009], a qual consiste em colocar os *fitness* normalizados dos indivíduos em uma roleta, a área a ser ocupada por cada indivíduo é proporcional ao valor de seu *fitness*, dessa forma os indivíduos com maior *fitness* tem mais chances de serem selecionados como pais.

d) Cruzamento (*crossover*): neste trabalho, foram utilizados os cruzamentos de dois pontos e uniforme. No cruzamento de dois pontos são sorteados os dois pontos de corte do par selecionado (cromossomos pais). No cruzamento uniforme uma máscara binária gerada de forma aleatória, determina de qual dos pais será selecionado o gene a ser preenchido nos cromossomos filhos. Na Figura 4, a máscara utilizada no cruzamento uniforme foi [0,0,1,0,0,1,1,1,0,0,1,0,1,0,1,0]. Quando o valor da máscara é zero, o gene a ser herdado pelo filho será do primeiro pai, e caso este valor seja um, herdará o gene do segundo pai. Devido ao fato de fazer um sorteio para cada posição, o cruzamento uniforme em média tem o seu desempenho superior ao cruzamento de dois pontos [Linden 2008].

e) Mutação: a cada 5 gerações ocorre a mutação.

f) Atualização: os indivíduos criados (filhos) em cada geração são inseridos na população.

g) Finalização: nesta etapa são verificadas se as condições para o encerramento da evolução foram atingidas. Neste trabalho foi adotado como critério de parada a quantidade de gerações.

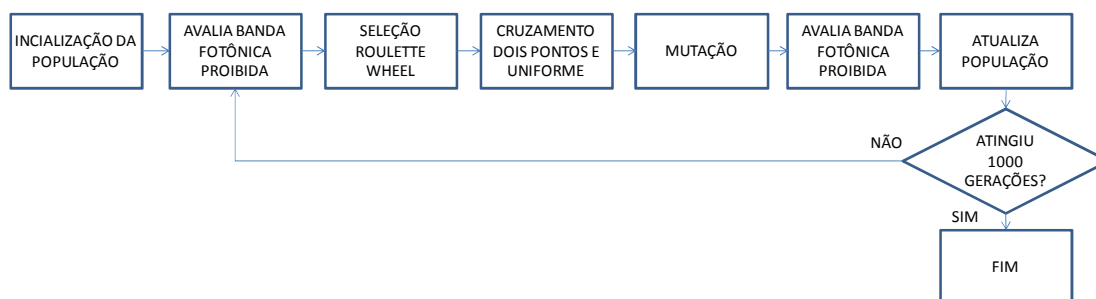


Figura 3. Fluxograma simplificado do Algoritmo Genético utilizado.

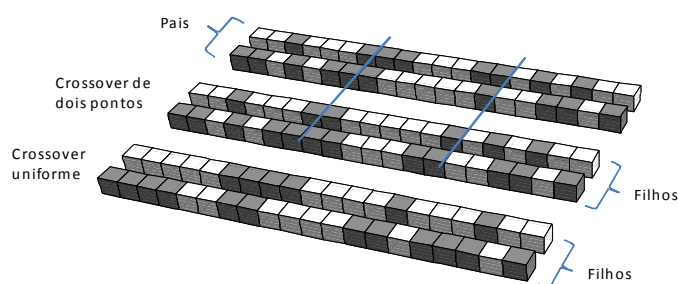


Figura 4. Cruzamento de dois pontos e o uniforme. As linhas representam os pontos de corte.

4. Cristais Fotônicos com Arranjos Quadrado e Triangular

Foram realizadas simulações com 20 e 50 indivíduos com elementos paralelogramos e triângulos, nas malhas da Figura 2, para os cruzamentos de dois pontos e uniforme. Um elemento paralelogramo é formado por dois triângulos adjacentes.

Nas Figuras 5 e 6 são apresentadas as estruturas cristalinas otimizadas para os arranjos quadrados e triangulares, respectivamente. Cada cristal é formado por 7 células unitárias nas direções horizontal e vertical. As estruturas obtidas com elementos triângulos foram similares às estruturas encontradas com elementos paralelogramos. Os resultados referentes à banda fotônica proibida dos cristais obtidos com diferentes cruzamentos e elementos estão apresentados na Tabela 1.

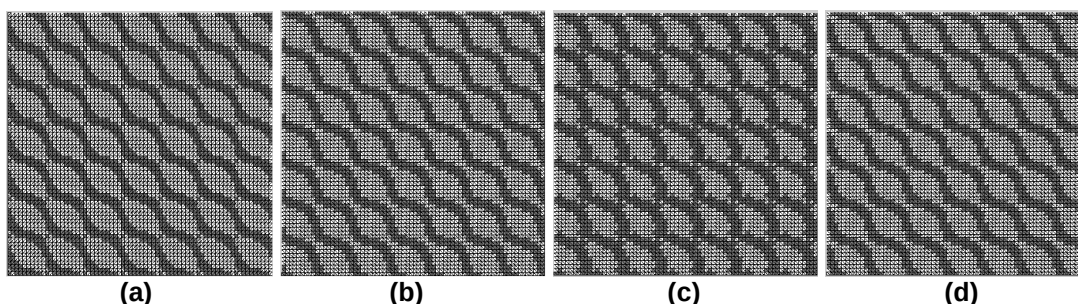


Figura 5. Cristais Fotônicos do modo TM_{12} para o arranjo quadrado com elementos paralelogramos (a) cruzamento de dois pontos com 20 indivíduos (b) cruzamento uniforme com 20 indivíduos (c) cruzamento de dois pontos com 50 indivíduos (d) cruzamento uniforme com 50 indivíduos.

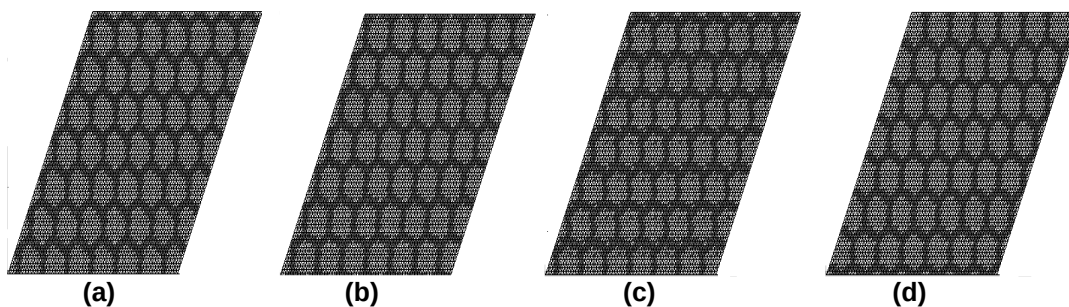


Figura 6. Cristais Fotônicos do modo TM_{12} para o arranjo triangular com elementos paralelogramos (a) cruzamento de dois pontos com 20 indivíduos (b) cruzamento uniforme com 20 indivíduos (c) cruzamento de dois pontos com 50 indivíduos (d) cruzamento uniforme com 50 indivíduos.

Com base nos resultados apresentados na Tabela 1, foi verificado que com o cruzamento uniforme obteve-se maiores Bandas Proibidas Relativas (ver Equação 1) quando o número de gerações é constante, pois este tipo de cruzamento explora um maior espaço de busca, já que os filhos gerados por esse cruzamento têm sua composição gene a gene, o que permite uma maior diversidade da população. O aumento da banda proibida relativa para o arranjo triangular foi mais acentuado, do que para o arranjo quadrado, o que demonstra que para este tipo de arranjo o cruzamento uniforme é o mais adequado.

Tabela 1. Banda Proibida Relativa para o Modo TM_{12}

Cristal com Arranjo Quadrado							
20 indivíduos				50 indivíduos			
Elemento Paralelogramo		Elemento Triângulo		Elemento Paralelogramo		Elemento Triângulo	
Dois Pontos	Uniforme	Dois Pontos	Uniforme	Dois Pontos	Uniforme	Dois Pontos	Uniforme
30,95%	31,67%	29,20%	34,32%	26,49%	31,37%	27,06%	34,25%
Cristal com Arranjo Triangular							
20 indivíduos				50 indivíduos			
Elemento Paralelogramo		Elemento Triângulo		Elemento Paralelogramo		Elemento Triângulo	
Dois Pontos	Uniforme	Dois Pontos	Uniforme	Dois Pontos	Uniforme	Dois Pontos	Uniforme
48,09%	49,18%	48,51%	51,74%	44,81%	49,20%	42,27%	51,65%

As Figuras 7 e 8 apresentam um comparativo da evolução do melhor indivíduo para cada tipo de cruzamento e tipo de elemento. Foi verificado que no cruzamento uniforme o valor ótimo foi atingido antes de completar as 1000 gerações, ou seja, convergiu mais rápido que o cruzamento de dois pontos.

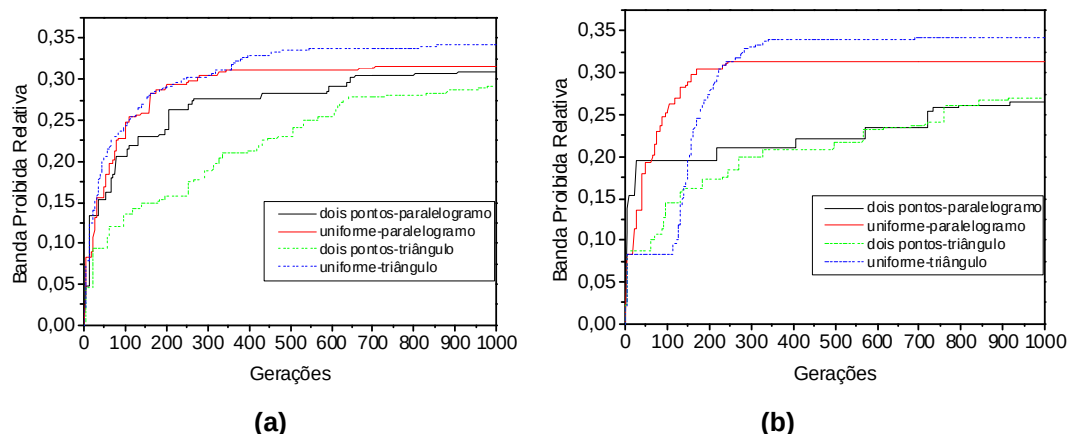


Figura 7. Evolução do melhor indivíduo do cristal com arranjo quadrado: (a) população com 20 indivíduos e (b) população com 50 indivíduos.

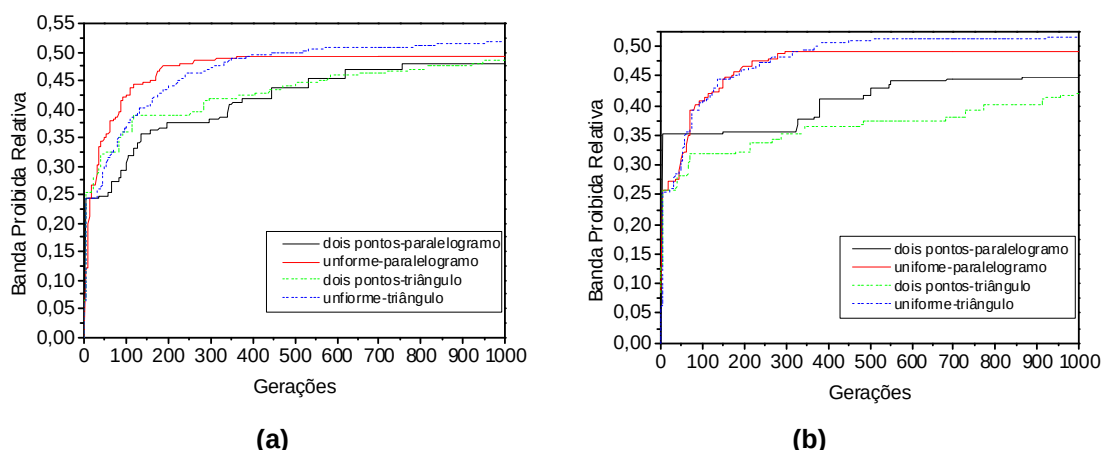


Figura 8. Evolução do melhor indivíduo do cristal com arranjo triangular: (a) população com 20 indivíduos e (b) população com 50 indivíduos.

A Tabela 2 apresenta o tempo médio de processamento de cada geração para ambos esquemas de cruzamentos, tipo de elementos e arranjos cristalinos. Para todas as simulações utilizou-se um Intel® Celeron® CPU 430 @ 1,80GHz com 2,0 Gb de RAM.

Tabela 2. Tempo médio de processamento para cada geração.

	Cristal com Arranjo Quadrado			
	Dois pontos		Uniforme	
	20 indivíduos	50 indivíduos	20 indivíduos	50 indivíduos
Elemento paralelogramo	1min 42s	3min 36s	1min 30s	3min 24s
Elemento triângulo	1min 12s	3min 24s	1min 18s	3min 12s
	Cristal com Arranjo Triangular			
	Dois pontos		Uniforme	
	20 indivíduos	50 indivíduos	20 indivíduos	50 indivíduos
Elemento paralelogramo	2min	5min	2min 18s	5min 18s
Elemento triângulo	2min	4min 56s	2min	5min

5. Conclusões

Com base nos resultados obtidos é possível afirmar que para o problema de otimização da banda fotônica proibida para o modo TM_{12} dos cristais fotônicos bidimensionais o cruzamento uniforme mostrou-se melhor que o cruzamento de dois pontos. Isto pode ser atribuído a maior diversidade dada à população com o cruzamento uniforme, o que permitiu ampliar o espaço de busca e localizar novos pontos de ótimos. Essa melhoria obtida com o cruzamento uniforme ficou mais evidente para o arranjo triangular, na qual banda fotônica relativa apresentou ganhos mais significativos em relação ao cruzamento de dois pontos.

O estudo com outros algoritmos bio-inspirados e a utilização de outros materiais na composição dos cristais fotônicos serão considerados em trabalhos futuros.

Os autores gostariam de agradecer o suporte financeiro da UFBA, IFBA, INCT Fotonicom, FAPESB, CAPES e CNPq Processo: 302390/2009-0.

Referências

- Affenzeller, H., Wagner, S., Winkler, S. e Beham, A. (2009) “Genetic Algorithms and Genetic Programming: Modern Concepts and Practical Applications.” Editora: CRC Press, Nova York-NY, páginas: 6-7.
- Haupt, R. L. e Werner, D. H. (2007) “Genetic Algorithms in Electromagnetics.” New York: Wiley.
- Joannopoulos, J. D., Johnson, S. G., Winn, J. N. e Meade, R. D. (2008) “Photonic Crystals: Molding the Flow of Light”, 2ª ed., Princeton University Press.
- Linden, R. (2008) “Algoritmos Genéticos: Uma importante Ferramenta da Inteligência Computacional”, Editora: Brasport, Rio de Janeiro-RJ.
- Malheiros, G. N. S., Esquerre, V. F. R. (2007) “Photonic Crystal Band Optimization by Genetic Algorithms”, Microwave and Optoelectronics Conference, 2007. IMOC 2007. SBMO/IEEE MTT-S International, páginas: 734-737.
- Malheiros, G. N. S., Esquerre, V. F. R. e Figueroa, H. E. H. (2011) “ A Strategy of Search e Refinement by GA in 2D Photonic Crystals with Absolute PBG”, Identificador: JQE-132416-2010, Número de páginas: 08. IEEE Journal of Quantum Electronics.
- Preble, S., Lipson, M., e Lipson, H. (2005) “2-D Photonic Crystals Designed by Evolutionary Algorithms.” Appl. Phys. Lett., vol. 86, nº. 6, pp. 061111(3).

Avaliação da Usabilidade em Sistemas de Gerenciamento de Biblioteca: Software SIGA e SophiA

Aparecida de Fátima Vaz¹, Maria Abadia Pires¹, Cristiane de Fátima dos Santos Cardoso¹

¹Instituto Federal Goiano Campus – Urutaí (IFG)
Rodovia Geraldo Silva Nascimento Km 2,5. CEP 75790-000 - Urutaí - Goiás - Brasil.
Fone/Fax: (64) 3465-1900

{cidinha_86v, maria_a_pires}@hotmail.com, crisfsantos@gmail.com

Abstract. *Usability is a key requirement for assessing the quality of software, verification of usability goals will indicate whether it will meet the user satisfaction or not. Therefore, this work studies were made beyond usability tests related to management systems and library SIGA Sophia. The usability tests include experienced users, intermediate and basic and are based on observation, interviews and performing tasks.*

Resumo. A usabilidade é um dos principais requisitos para a avaliação da qualidade de um software, a verificação das metas de usabilidade irá indicar se o mesmo atenderá o usuário satisfatoriamente, ou não. Por isso, neste trabalho foram feitos estudos sobre usabilidade além de testes referentes aos Sistemas de gerenciamento de biblioteca SIGA e SophiA. Os testes de usabilidade contemplam usuários experientes, intermediários e básicos e são baseados em observação, entrevistas e realização de tarefas.

1. Introdução

No século XX com a expansão do uso de computadores surgiram inúmeras pesquisas em IHC (Interação Humano Computador) e ergonomia, mas foi em meados da década de 80 que a palavra usabilidade ganhou notoriedade, ela passou a ser usada para caracterizar o trabalho do usuário ao utilizar o computador. Rebelo (2009) traz a definição de usabilidade criada por Nielsen, que estabelece “a usabilidade é um atributo de qualidade que avalia quão fácil uma interface é de usar”, ou seja, reflete a qualidade da interação com um sistema e está relacionada a diretrizes e metas que facilitem a vida do usuário.

Diversos são os fatores que podem influenciar na interação com o software, até mesmo o contexto no qual está inserido, em geral o usuário está envolvido em diversas tarefas simultaneamente e isso pode ser prejudicial ao bom uso do software. Em bibliotecas esta situação é bastante característica, nelas diversas tarefas são realizadas em um curto intervalo de tempo, e devem ser feitas com precisão e eficiência. As ações de consultar, emprestar e devolver obras devem ser realizadas com facilidade e rapidez, assim o trabalho do bibliotecário terá maior qualidade no tratamento da informação, com isso os usuários ficarão mais satisfeitos e a universidade ganha, facilitando no processo de disseminação do conhecimento. Por isso ter um software que gerencie as

principais funções de forma amigável torna-se imprescindível, afinal, a biblioteca universitária é um dos pontos de referência em uma universidade, representando um elemento dinâmico e propulsor trazendo contribuição científica, técnica, cultural, humanística e didática.

Portanto, é de grande valia uma análise da usabilidade de sistemas gerenciadores de biblioteca. A seguir serão apresentados conceitos sobre usabilidade, metas decorrentes da experiência do usuário, e os resultados dos testes de usabilidade em dois softwares: SIGA e Sophia, sendo o primeiro livre e o segundo proprietário.

2. Usabilidade

De acordo com Nielsen (1993, p. 26, apud Costa e Ramalho, 2010, p. 111) a usabilidade está associada a vários fatores, tais como, a facilidade de aprendizado, eficiência de uso, facilidade de memorização, baixa taxa de erros e satisfação subjetiva. Em conjunto, estes fatores constituem os quesitos primordiais para sua mensuração e consequentemente a qualidade do software.

Compreender a usabilidade é um processo inerente ao bom desempenho do software, as metas de usabilidade são utilizadas como forma de verificar problemas quanto a produtividade do usuário ao utilizar o sistema e ao aprendizado eficaz, elas também fornecem soluções de interação. Com isso, Preece (2005) define seis metas de usabilidade: **eficácia** - é uma meta bastante geral e se refere ao quanto um sistema é bom em se fazer o que se espera dele; **eficiência** se refere à maneira como o sistema auxilia os usuários na realização de suas tarefas; **segurança** implica em proteger o usuário de condições perigosas e situações indesejáveis; **utilidade** refere-se à medida na qual o sistema propicia o tipo certo de funcionalidade, de maneira que os usuários possam realizar aquilo que precisam ou que desejam; **capacidade de aprendizagem (learnability)** refere-se a quão fácil é aprender a usar o sistema; **capacidade de memorização (memorability)** refere-se à facilidade de lembrar como utilizar um sistema, depois de já ter aprendido como fazê-lo – algo especialmente importante para sistemas interativos que não são utilizados com muita frequência.

A usabilidade é considerada negativa ou positiva. A usabilidade negativa ocorre quando existe uma má utilização do sistema, contribuindo para sua ineficiência e abandono. Quanto à positiva denota as relações interpessoais, comerciais e informacionais necessitadas pelo usuário, satisfazendo plenamente a interação com o sistema.

3. Metas Decorrentes da Experiência do Usuário

Utilizar apenas as metas de usabilidade não é suficiente para determinar a qualidade de um sistema, junto destas há um componente pessoal e subjetivo que deve ser levado em consideração. Preece et al. (2005) fala das metas decorrentes da experiência do usuário, que foram criadas para complementar a noção de quão boa é a interação com o software. Portanto, os sistemas devem ser concebidos para serem satisfatórios, agradáveis, interessantes, úteis, motivadores e compensadores, aumentando assim a eficiência e produtividade da interação.

Preece et al. (2005) ressalta ainda que, equilibrar as metas de usabilidade e as decorrentes da experiência do usuário é imprescindível, e lembra que nem todas as metas se aplicam a todo produto, pois algumas combinações são incompatíveis.

4. Sistemas gerenciadores de biblioteca

O uso de TI (Tecnologia da Informação) nas bibliotecas vem se aprimorando com o passar dos anos, existem vários sistemas gerenciadores de biblioteca no mercado, alguns gratuitos e outros não. Neste trabalho dois softwares foram testados, SophiA e SIGA, ambos constituídos de duas partes: o sistema central – ou administrativo, no qual administradores interagem e o módulo online direcionado aos usuários da biblioteca.

SophiA Biblioteca é um software de gerenciamento de biblioteca proprietário muito utilizado, são mais de 2500 clientes, é um sistema de solução modular que opera de acordo com as normas e padrões da biblioteconomia, mas sem exigir conhecimento profundo em informática do bibliotecário. O sistema é oferecido em três versões distintas (Básico, Intermediário e Avançado), sendo que cada versão é composta pelo Módulo Gerenciador e por módulos opcionais que permitem ao usuário expandir o potencial do sistema, o SophiA permite a informatização da biblioteca de acordo com as necessidades da Instituição, maiores informações podem ser obtidas em (http://www.primasoft.com.br/2006/html/interna_1.php?cod=33).

Um outro software para gerenciamento de biblioteca é o Sistema Integrado de Gestão Acadêmica (SIGA), que têm como base para seu desenvolvimento o framework Miolo, que são aplicações acessíveis via web, utilizando orientação a objetos através da linguagem PHP5. O sistema SIGA possui módulos que cobrem diferentes áreas, sendo as principais: a biblioteca, recursos humanos e administração, lembrando que para o desenvolvimento deste trabalho apenas o módulo de Biblioteca é avaliado. O SIGA é um software livre e faz parte de um grande projeto do governo para informatização das instituições federais de ensino, maiores informações podem ser obtidas em: (http://www.softwarepublico.gov.br/ver-comunidade?community_id=20428249).

5. Revisão Literária

Na literatura não se encontra muitos trabalhos semelhantes a este. Um trabalho que pode ser destacado é apresentado por Evaristo (2011), que analisa as telas de interface do sistema SophiA sob a ótica do bibliotecário, a usabilidade da tarefa de catalogação do sistema é avaliada com base na opinião de 5 usuários e o método de análise é baseado em questionários para medir o grau de satisfação quanto aos itens de interface/interação.

Já Marcelino (2008) apresenta uma análise sobre a usabilidade na biblioteca online do INPE (Instituto Nacional de Pesquisas Espaciais) na qual são realizados testes com 10 usuários através de anotações e gravações em vídeo afim de identificar os problemas de usabilidade e fazer sugestões para melhoria. No referido trabalho os aspectos visados são: a facilidade de aprendizagem, a baixa taxa de erros, eficiência e satisfação subjetiva.

Tendo em vista as diversas maneiras de avaliar a usabilidade Pereira (2011), avaliou a interface por meio de um método analítico (avaliação heurística) e um método empírico (teste com usuários), obtendo resultados tanto dos avaliadores quanto dos usuários. Diante disso, o presente trabalho utiliza métodos empíricos com uma grande quantidade de usuários e cada meta de usabilidade é avaliada, verifica-se também as o alcance das expectativas do usuário através das metas decorrentes da experiência do usuário.

6. Experimento - Testes de Usabilidade

Os testes de usabilidade dos softwares SIGA e SophiA foram realizados no Instituto Federal Goiano – *Campus* Urutaí, contemplando 50 usuários dos mais diversos: experientes, intermediários e básicos nos módulos online, administradores e avançados na parte administrativa do sistema. Para os módulos online foram escolhidos usuários com pouquíssimo ou nenhum conhecimento dos sistemas em questão, sendo a sua categorização realizada em função da experiência no uso de computadores. Já nos módulos administrativos a classificação foi realizada com base no conhecimento prévio, usuários frequentes do SophiA foram classificados como “administrador”, e usuários experientes, sem conhecimento no uso de software gerenciador de biblioteca foram classificados como “avançado”. A princípio seriam feitos testes também com usuários frequentes do SIGA, mas como este sistema ainda está em implantação os testes não foram possíveis, portanto, espera-se algumas distorções a favor do SophiA nos módulos administrativos, essa questão será tratada nas análises.

Os testes consistem da realização de tarefas por parte de tais usuários e aplicação de questionários com o intuito de fazer um comparativo entre os dois softwares. A partir dos resultados obtidos é possível verificar se as metas de usabilidade estão sendo cumpridas. Foi estabelecido um roteiro de atividades para o teste, e na realização do mesmo, as tarefas realizadas foram cronometradas para definir a eficiência e eficácia. Para medir o aprendizado e a memorização foi necessário dividir o teste em duas etapas: na primeira as tarefas foram realizadas com o auxílio dos avaliadores e na segunda os usuários realizaram sozinhos. Foi realizado um pré-teste com usuários selecionados, apenas 4, para estabelecer medidas de tempos aceitáveis para os testes subsequentes, além disso, foi determinado o número de usuários para estes testes, os valores podem ser vistos na tabela 1.

Tabela 1. Dados quantitativos para realização das tarefas

Software/ Módulo	Tempo Estimado	Tempo de eficiência	Número de usuários	Especificação dos usuários
SophiA	8 min e 30 segundos	6 min e 48 segundos	10	5 administrativos e 5 avançados
SophiA on- line	3 minutos	2min e 24 segundos	15	5 experientes 5 intermediários 5 básicos
SIGA	8 min e 30 segundos	6 min e 48 segundos	10	5 administrativos 5 avançados
SIGA on-line	3 minutos	2min e 24 segundos	15	5 experientes 5 intermediários 5 básicos

Os princípios avaliados a partir da aplicação do teste foram: eficácia (ser eficaz no uso), eficiência (ser eficiente no uso), segurança (ser seguro no uso), utilidade (ser de boa utilidade), learnability (ser fácil de aprender), memorability (ser fácil de lembrar como se usa). E ainda as metas decorrentes da experiência do usuário: ser satisfatório, agradável, interessante, útil, motivador e esteticamente apreciável. Os resultados, que podem ser vistos nas tabelas 2 e 3, e na figura 1, serão discutidos a seguir.

5.1 Eficácia e eficiência

A eficácia pode ser entendida como o alcance dos objetivos, enquanto que a eficiência de um sistema está relacionada com a velocidade de uso do usuário para alcançar seus

objetivos. Com base nessas informações, as seguintes métricas foram definidas:

Taxa de eficácia - porcentagem de usuários que conseguiram terminar as tarefas dentro do tempo determinado.

Taxa de eficiência - porcentagem de usuários que conseguiram terminar as tarefas em tempo igual ou inferior ao tempo determinado para eficácia reduzido em 20%.

Como pode ser visto na tabela 2, em termos de eficácia dos módulos online o desempenho foi ligeiramente melhor no SIGA, os usuários intermediários conseguiram rapidamente compreender e interagir com o sistema. Já a eficiência foi idêntica, com a diferença que apenas usuários experientes do SophiA conseguiram realizar as tarefas no tempo determinado, enquanto no SIGA online um usuário intermediário conseguiu. Contudo, o que chama a atenção é o desempenho na parte administrativa, em que a diferença foi bastante acentuada, era esperado um excelente desempenho para o SophiA em função do conhecimento prévio, o que de fato ocorreu. Os usuários avançados que não possuíam conhecimento prévio em nenhum dos sistemas foram mais eficazes no SIGA - foram em média 6% mais rápidos em relação aos administradores e no SophiA em média 25% mais lentos. No tocante à eficiência, o SIGA deixou bastante a desejar.

Assim, as conclusões que se pode tirar é que o SIGA é mais compreensível para os novos usuários, mas não os ajuda a ter um desempenho melhor em interações subsequentes. Também há de se considerar os fatores comportamentais que se traduzem como resistência a novos paradigmas. Conforme é dito por Turbam et al. (2007), em face de novos sistemas ocorre uma tendência natural do usuário em relutar e optar pelos métodos antigos, (TURBAN ET AL., 2007).

Tabela 2. Dados sobre taxa de eficiência, eficácia e aprendizado/memorização

Software/ Módulo	Tipos de Usuário	Taxa de Eficácia	Taxa de Eficiência	Percentual de melhora do tempo entre uma rodada de interação e outra (aprendizado/memorização)
Sistema SophiA	Administrador	100%	60%	28,10%
	Avançado	60%	40%	34,76%
Sistema SophiA on-line	Básico	0%	0%	48,96%
	Intermediário	40%	20%	54,63%
	Experiente	100%	100%	43,02%
Sistema SIGA	Administrador	40%	20%	41,72%
	Avançado	80%	0%	34,79%
Sistema SIGA on-line	Básico	0%	0%	43,60%
	Intermediário	80%	20%	33,64%
	Experiente	100%	80%	30,97%

5.2 Análise da Segurança e da utilidade

A segurança serve para a prevenção de erros e redução de riscos, como por exemplo as proximidades de botões antagônicos salvar-excluir. Em relação à segurança do software, como pode ser visto na figura 1, o SophiA obteve um melhor resultado que o SIGA, as principais reclamações foram quanto a falta de organização dos ícones e a ausência de botões para retorno ou cancelamento de algumas ações. O SophiA também apresentou como ponto forte o *feedback*, que deixa o usuário bastante seguro. Nos módulos online a avaliação foi a mesma.

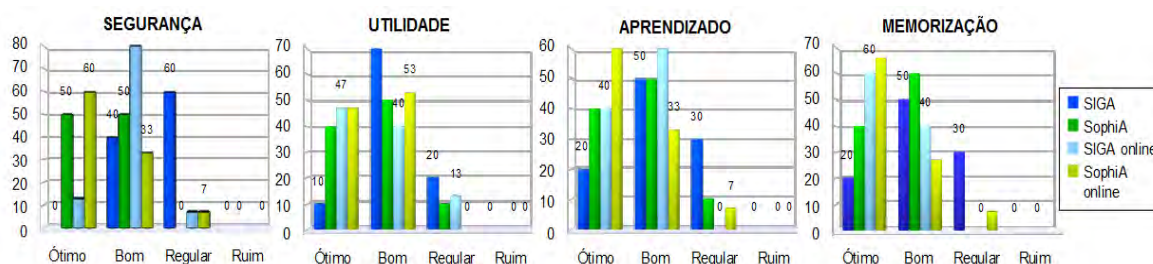


Figura 1 – Gráficos das metas: segurança, utilidade, aprendizado e memorização

Para que um sistema tenha utilidade ele deve fornecer o tipo certo de funcionalidade que o usuário espera dele. No sistema SIGA foram relatados problemas como: dificuldades na localização dos ícones na interface e também dispersão de alguns formulários. Já o sistema SophiA teve problemas com o excesso de campos e má localização dos ícones principais. Com isso, o SophiA administrativo foi melhor avaliado, mesmo pelos usuários sem conhecimento prévio, enquanto que os módulos online tiveram desempenho semelhante. Poucos usuários consideraram o SIGA administrativo ótimo no quesito utilidade.

5.3 Análise do Aprendizado e Memorização

O aprendizado também é conhecido como *Learnability* e determina a facilidade de aprender a utilizar um sistema. Como pode ser visto na figura 1, cerca de 30% dos usuários do sistema SIGA o classificaram como regular, pois tiveram dificuldade no aprendizado em função do excesso de ícones e vocabulários incomum. Já os usuários que o classificaram como bom ou ótimo consideraram os ícones como fáceis, ou seja, com a presença de metáforas. Curiosamente o módulo online do SIGA foi o único que não obteve voto de regular.

A memorização está bastante relacionada ao aprendizado, também é conhecida como *memorability* e é descrita quando existe a facilidade de lembrar como utilizar o um sistema após uma experiência prévia. O sistema SophiA recebeu maior aprovação dos usuários quanto ao conceito ótimo tendo 40%, e para o SIGA 20%. Contudo, 30% dos usuários do SIGA o classificaram como regular, pois encontraram dificuldades na memorização devido a ocultação de ícones, dificultando a pesquisa bem como a ausência de botões cancelar e voltar.

Tanto a memorização quanto o aprendizado estão intrinsecamente ligados a experiências prévias. Por isso, os testes foram realizados duas vezes e o tempo medido. Em seguida foi feito o cálculo do quanto o usuário melhorou entre uma rodada de interação e outra, foram obtidos os seguintes valores: no SophiA os usuário apresentaram em média 31,43% de melhora nos tempos, no módulo online a melhora foi bastante significativa e ficou em torno de 48,83%. Já no SIGA a melhora foi de 38,25% e no módulo online de 32,31%. O que claramente evidencia a superioridade do SophiA online sobre os demais em termos de aprendizado e memorização. É importante observar que estes valores englobam todos os indivíduos envolvidos nos testes, a tabela 2 mostra o índice de melhora dos tempos por categoria de usuário.

6. Resultados das Metas decorrentes da experiência do usuário

Como foi dito, a satisfação do usuário é parte importante do projeto e tem por objetivo oferecer mais qualidade ao produto, assim, segue a avaliação de algumas metas, os números relativos a esta avaliação podem ser vistos na tabela 3.

Para que um software seja satisfatório, ele deve desempenhar e corresponder todas as expectativas dos usuários. Assim, na tabela 3 observa-se a superioridade do sistema SophiA sobre o SIGA na meta “ser satisfatório”. E para que o sistema seja agradável ele deve proporcionar uma interação transparente em que o usuário não se sintá impedido ou impossibilitado pelo sistema, pode ser percebido que o sistema SophiA consegue agradar mais os usuários, tanto no sistema como no módulo online.

Tabela 3: Metas decorrentes da experiência do usuário

Meta “ser satisfatório”				
Módulo	Ótimo	Bom	Regular	Ruim
SIGA	10%	70%	10%	10%
SophiA	40%	50%	10%	-
SIGA online	47%	40%	13%	-
SophiA online	40%	60%	-	-
Meta “ser agradável”				
SIGA	-	70%	30%	-
SophiA	40%	40%	20%	-
SIGA online	20%	73%	7%	-
SophiA online	47%	53%	-	-
Meta “ser interessante”				
SIGA	20%	50%	30%	-
SophiA	40%	30%	30%	-
SIGA online	47%	30%	13%	-
SophiA online	67%	20%	13%	-
Meta “ser útil”				
SIGA	20%	70%	-	10%
SophiA	70%	30%	-	-
SIGA online	53%	40%	7%	-
SophiA online	67%	33%	-	-
Meta “ser esteticamente apreciável”				
SIGA	10%	40%	30%	20%
SophiA	10%	50%	30%	10%
SIGA online	13%	74%	13%	-
SophiA online	39%	27%	27%	7%
Meta “ser motivador”				
SIGA	10%	30%	50%	10%
SophiA	50%	20%	30%	-
SIGA online	27%	53%	20%	-
SophiA online	40%	33%	27%	-

O sistema sendo interessante para o usuário provoca mais curiosidade, correspondendo assim, suas expectativas ao utilizá-lo. Nesta meta destaca-se o SophiA online, conforme pode ser observado na tabela 3 em que 67% dos usuários o consideram ótimo. O sistema também deve ser proveitoso e vantajoso, facilitando seu uso pelos usuários. Observa-se que o sistema SophiA, no quesito utilidade, obteve uma

porcentagem alta em relação ao SIGA, 70% contra 20% de conceito ótimo. Neste caso prevaleceu também a experiência prévia dos usuários administradores.

O sistema é motivador quando ele desperta interesse do usuário e estimula a interação com o software. Essa meta chama bastante a atenção, pois o desempenho do SIGA foi bastante inferior. A diferença se dá, pelo fato de que, ao abrir o sistema SIGA o usuário se depara com vários ícones espalhados na interface, o que dificulta as ações, em contrapartida, no SophiA a interface é mais organizada, facilitando a visualização dos ícones.

O tipo de busca realizada pelo SIGA prejudicou bastante sua avaliação no geral, o mesmo não aceita qualquer tipo de acentuação nas palavras e não possui a opção de alto complemento, sendo necessária a digitação correta das palavras.

7. Conclusão

Em função da importância de se verificar a usabilidade de um sistema, este trabalho avaliou dois sistemas gerenciadores de biblioteca, e como resultado desse experimento pode-se perceber que dentre os softwares avaliados o SophiA teve um desempenho ligeiramente melhor. O principal ponto que levou a aprovação do software foi o fato dele possuir as características desejáveis em um software desta categoria: a segurança, a motivação e a utilidade. Por outro lado, os usuários tiveram dificuldades quanto à localização de funções, pois o sistema possui várias metáforas, as quais nem sempre estão familiarizados, e as fontes são pequenas dificultando a visualização. O sistema SIGA em alguns quesitos como a memorização não teve muita diferença em relação ao SophiA, porém os usuários tiveram algumas dificuldades quanto a localização dos ícones na interface, dispersão de formulários e operações de busca.

Como trabalho futuro pretende-se fazer uma análise mais detalhada de algumas funcionalidades do software SIGA e aplicar outros métodos de avaliação de usabilidade.

Referências bibliográficas

- Costa, L. F. e Ramalho, F.A. (2010) “A Usabilidade nos Estudos de Uso da Informação: em cena usuários e sistemas interativos de informação”. *Perspectivas em ciência da informação*, v.15, n.1, p.92-117, jan./abr.
- Evaristo, J. B. (2011) “Usabilidade da Tarefa de Catalogação no Sistema de Automação SophiA biblioteca”, Monografia, Universidade Federal de Brasília.
- Marcelino, S. C. O (2008) “Estudo de usuários de usabilidade de sites de bibliotecas especializada: o caso da 'biblioteca on line' do Instituto Nacional de Pesquisas Espaciais (INPE)”. Monografia, Instituto Nacional de Pesquisas Espaciais.
- Pereira, F. O (2011) “Avaliação de Usabilidade em Bibliotecas Digitais: Um estudo de caso.” Monografia, Universidade Federal de Minas Gerais.
- Preece, J. (2005) “Design de Interação: Além da Interação Homem-Computador”. Porto Alegre: Bookmam.
- Rebelo, I. B. (2009) “Interação entre Homem e Computador”. *Tecnologias Aplicadas a Sistemas de Informação – TASI Apostila desenvolvida para o curso de Sistemas de Informação*. Centro Universitário Unieuro.
- Turban, E. Rainer Jr, R. K. e Potter, R. (2007) “Introdução a sistemas de informação: uma abordagem gerencial”. Rio de janeiro: Elsevier.

Padrões e diretrizes arquiteturais para escalabilidade

Acrísio J. Nascimento Jr¹, Pedro Frosi Rosa², Ivens Oliveira Porto², Michel dos Santos Soares²

¹Ciência da Computação – Universidade Federal de Goiás/Campus Catalão (UFG/CAC)
Caixa Postal 536 – 75.704-020 – Catalão – GO – Brazil.

²FACOM – Universidade Federal de Uberlândia (UFU) – Uberlândia, MG – Brazil.
acrisiojr@gmail.com, {frosi,michel}@facom.ufu.br,
ivens.porto@gmail.com

Abstract. *One issue not fully explored is how to build an architecture for a scalable system. There are works that discuss principles and general techniques for scalability, specially about performance improvement. However, this information is disorganized and unstructured, and are, to an architect with the responsibility of building a scalable system, a poor source of information to support his work.*

Resumo. *Uma questão não totalmente explorada é como construir uma arquitetura para um sistema escalável. Há obras que discutem princípios e técnicas gerais para a escalabilidade, especialmente sobre melhoria de desempenho. No entanto, esta informação é desorganizada e não estruturada, e são, para um arquiteto com a responsabilidade de construir um sistema escalável, uma fonte pobre de informação para apoiar o seu trabalho.*

1. Introdução

Nos dias atuais, escalabilidade tornou-se uma importante propriedade em sistemas. Com o volume crescente de usuários e dados que trafegam nas redes, os sistemas devem atender a esta demanda de maneira mais eficiente possível. Então, escalabilidade tornou-se um desafio para os arquitetos e desenvolvedores. Um exemplo de uso mais evidente em relação a esta propriedade são as ferramentas de busca (dentre elas, a mais popular: Google). Então, como arquitetar um sistema escalável?

O objetivo deste artigo é o de identificar, discutir as diretrizes e técnicas arquiteturais para auxiliar arquitetos de sistemas a projetar e construir sistemas escaláveis. As diretrizes e padrões deste trabalho são aplicáveis particularmente a sistemas web e sistemas de rede que trabalham com dados armazenados. Durante a fase de projeto, devem-se observar as diretrizes e técnicas de escalabilidade para que não se cometa erros no desenvolvimento. Não existe uma estratégia ou diretriz que solucione todos os problemas relacionados à escalabilidade.

Neste trabalho, as diretrizes e técnicas se preocupam principalmente com a escalabilidade horizontal, tornando possível a execução de um sistema em vários nós de processamento de tal maneira que, ao aumentar a quantidade de nós o sistema aumente, ou mantenha, seu desempenho de maneira satisfatória.

A escalabilidade horizontal tem um fator positivo, custo, que auxilia bastante no momento de determinar qual tipo de escalabilidade optar. Além da limitação a que um determinado nó se encontra para se escalar verticalmente.

A escalabilidade vertical possui algumas restrições, uma delas o custo, que inviabiliza seu uso se comparado com as demais formas de escalabilidade. Outro ponto importante é a complexidade de software em uma escalabilidade vertical.

2. Posicionamento de contexto em escalabilidade de sistemas

2.1. Categorias de Escalabilidade

Escalabilidade linear: significa que para cada recurso adicionado a um sistema, o desempenho aumenta de maneira diretamente proporcional, como mostra a Figura 1. O fator de escalabilidade aqui é igual a 1 [Williams, 2004].

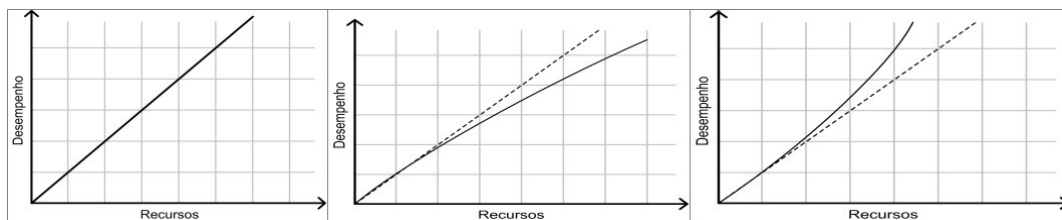


Figura 1. Escalabilidade linear, sublinear e superlinear

Escalabilidade sublinear: significa que para cada recurso adicionado a um sistema, o desempenho aumenta de maneira não proporcional, e inferior, à capacidade individual dos recursos adicionados, como mostra a Figura 1. O fator de escalabilidade é menor que 1 [Williams, 2004].

Escalabilidade super linear: como mostra a Figura 1 [Williams, 2004], significa que para cada recurso adicionado a um sistema, o desempenho aumenta de maneira superior à capacidade do recurso. O fator de escalabilidade é maior que 1. Isto se dá, porque quando se adiciona um determinado recurso, junto a este pode estar adicionado-se outros recursos.

2.3. Contexto de Escalabilidade

Construir sistemas de alto desempenho é diferente de se construir sistemas escaláveis. Para a construção de sistemas de alto desempenho aplicam-se técnicas para aumentar a velocidade de processamento e minimizar o uso de recursos pelo sistema, preocupando-se apenas em melhorar as métricas de desempenho de cada nó de processamento individual e não se considera a extensão do sistema como meio para aumentar o desempenho, como descrito na definição de escalabilidade horizontal.

No ano de 2000, Eric Brewer [Brewer, 2000] fez uma conjectura, que ficou conhecida como Teorema CAP, que diz o seguinte: “Dadas as propriedades de Consistência, Disponibilidade e Tolerância a Particionamento da rede, um serviço Web pode ter no máximo duas destas propriedades. Ela não se aplica apenas a serviços Web.

Consistência – um serviço é dito consistente se todos os clientes que acessam o objeto de dados vêem o mesmo dado, mesmo com a ocorrência de escritas concorrentes no objeto de dados.

Disponibilidade – para que um sistema distribuído esteja continuamente disponível, toda requisição recebida por um nó do sistema, que não apresente falhas, deve ter uma resposta. Em um contexto de acesso a dados, disponibilidade quer dizer que sempre será possível acessar o dado, mesmo que seja uma imagem (réplica) [Porto, 2009].

Tolerância a partições – as definições anteriores de Consistência e Disponibilidade são qualificadas com a necessidade de Tolerância a partições de rede [Porto, 2009].

A escolha das propriedades de Consistência e Disponibilidade resulta em sistemas onde é preciso que todos os nós devem comunicar-se para manter a consistência dos dados. Exemplos de sistemas deste tipo são: Bancos de Dados hospedados em um único local; e Bancos de Dados em *cluster* (ou qualquer outro tipo de arranjo). Algumas características marcantes destes sistemas são o uso de protocolo 2PC (*two phase commit*) e protocolos de invalidação de cache. Estes são sistemas distribuídos clássicos.

A escolha das propriedades de Consistência e Tolerância a Partições resulta em sistemas onde deve-se tolerar o fato de o sistema parar de responder quando ocorre um particionamento de rede, já que é preciso comunicação entre todos os nós para manter a consistência. Exemplos de sistemas deste tipo são Bancos de Dados distribuídos, Sistemas de Travas Distribuídas de dados. Algumas características marcantes destes sistemas são algoritmos de travas pessimistas, uso de protocolos de quorum, e o fato de que partições pequenas tornam o sistema indisponível.

A escolha das propriedades de Disponibilidade e Tolerância a Partições resulta em sistemas onde nem sempre se trabalha com dados atualizados. Exemplos de sistemas deste tipo são DNS (Domain Name System), caches WEB, Coda, Bayou [Demers et. al. 1994]. Suas características mais marcantes são uso de consistência temporal (TTL - time to live), uso de leases [Gray e Cheriton, 1989], algoritmos de travas otimistas e atualização otimista dos dados com possível resolução de conflitos.

3. Padrões Arquiteturais de Escalabilidade

Padrões de projeto (Design Pattern), em sua maioria, tratam de problemas relacionados a armazenamento de dados e uso de transações.

3.1. Definições e Aspectos

Elaborar arquiteturas de software não é uma tarefa fácil. Arquitetos e projetistas experientes quase nunca criam uma nova solução. Qualquer sistema é composto de subsistemas. Estes subsistemas disponibilizam uma interface para que possam ser utilizados por outros subsistemas. Quando subsistemas se tornam grandes e complexos, é interessante que se crie uma “fachada” de acesso ao subsistema, de tal modo que esta “fachada” unifique e facilite o uso do subsistema, como mostra a Figura 2 [Porto, 2009].

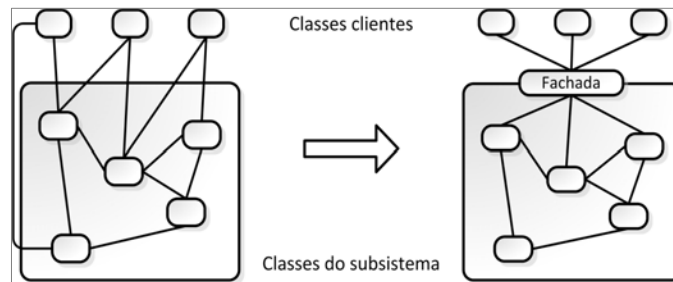


Figura 2. Padrão “fachada”

Uma definição ampla do que é um padrão, segundo [Porto, 2009], é a de que um padrão endereça um problema de projeto recorrente que surge em situações de projeto específicas, e apresenta uma solução. Padrão arquitetural: expressa o esquema organizacional estrutural fundamental de um sistema. Ele provê um conjunto predefinido de subsistemas, especifica suas responsabilidades, e inclui regras e diretrizes para organizar seus relacionamentos.

Um padrão de projeto é um esquema para refinar os componentes ou subsistemas de um sistema, ou o relacionamento entre eles. Ele descreve uma estrutura comumente recorrente de componentes que se comunicam para solucionar um problema genérico de projeto em um contexto particular [Porto, 2009], [Gamma et. al., 1995].

3.2. Vantagens no uso de padrões

Padrões documentam a experiência em projetos de software existentes, o que permite o uso sistemático da experiência adquirida em muitos anos de trabalho e auxilia arquitetos e projetistas menos experientes na criação de sistemas de melhor qualidade.

Identificam e especificam abstrações que estão acima do nível de classes individuais e instâncias de objetos e de componentes, e detalha suas responsabilidades e relacionamentos, sendo que estes componentes, juntos e em acordo como padrão, solucionam o problema colocado.

Estabelecem um vocabulário e um entendimento comuns dos princípios de projeto, onde os nomes dos padrões se tornam parte de uma linguagem de padrões que é compartilhada por todos e, esta linguagem facilita a comunicação, o entendimento do funcionamento do sistema, e a discussão de problemas de projeto, sendo que o nome do padrão pode abstrair sua complexidade e apenas com o nome ser possível comunicar toda a solução de um problema.

Auxiliam na construção de sistemas com propriedades definidas, além de tratar também de requisitos não funcionais, tais como escalabilidade, confiabilidade, reusabilidade, disponibilidade, etc.

3.3. Arquitetura Shared Nothing

Auxilia na construção de sistemas facilmente escaláveis na horizontal, a partir da estruturação do sistema em partes independentes, sem compartilhamento de estado (*Shared Nothing*), possibilitando a escalabilidade linear. Por exemplo, suponha um sítio onde o usuário forneça seus dados para cadastro e após isto, o mesmo seja autenticado na página para usufruir dos benefícios oferecidos pelo sítio. Normalmente, este tipo de

sistema é desenvolvido utilizando-se de uma arquitetura de 3 (três) camadas tradicional, sendo elas: apresentação, negócios e acesso a dados [Porto, 2009].

No intuito de aumentar o desempenho da solução, as camadas são fisicamente separadas, onde cada camada é composta por um *cluster* de servidores. As camadas de apresentação e de negócios têm seus dados relativos ao estado do uso do sistema pelo usuário armazenado em memória, como os que são utilizados por várias requisições.

Os servidores de cada camada se comunicam para sincronizar o trabalho e para aumentar a disponibilidade do sistema, além da comunicação entre as camadas para a sincronização de estado. À medida que uso do sistema cresce, adicionam-se novos servidores para atender à demanda, escalando-o horizontalmente.

A Figura 3 mostra o aumento da complexidade do sistema com a adição de novos servidores. Um modelo formal que descreve a escalabilidade linear é a lei de Amdahl [Amdahl, 1967].

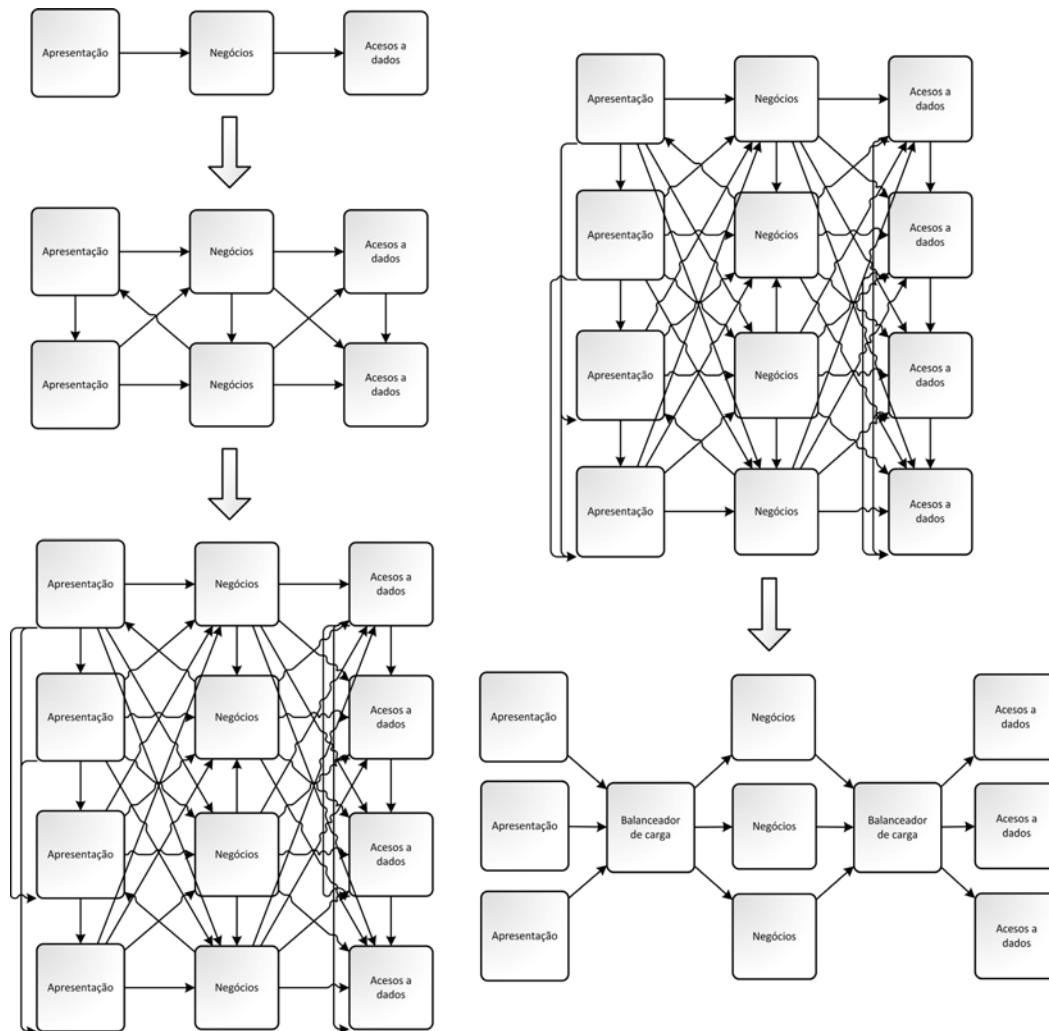


Figura 3. Aumento da complexidade do sistema com a adição de novos servidores e também uma possível solução com o uso de *Shared Nothing*

Podem haver sobretaxas de comunicação e sincronização, mas devem ser as menores possíveis. A idéia geral, em um sistema totalmente *Shared Nothing*, é remover toda a dependência entre os nós, como mostra a Figura 3.

A estrutura de uma Arquitetura *Shared Nothing* (SNA) é simples, basta que os nós do sistema sejam independentes. O resultado final da aplicação da SNA depende muito do problema, apesar do resultado ser sempre nós independentes, os nós resultantes e suas responsabilidades. A comunicação entre os nós de uma mesma camada não mais existirá, pois ter-se-á apenas comunicação entre camadas.

O padrão *Sharding*, mostrado na Figura 4, proporciona escalabilidade horizontal para os dados de um sistema, através do particionamento dos dados em vários bancos de dados, possibilitando uma alternativa de custo mais baixo e sem as limitações da escalabilidade vertical que normalmente é aplicada a banco de dados.

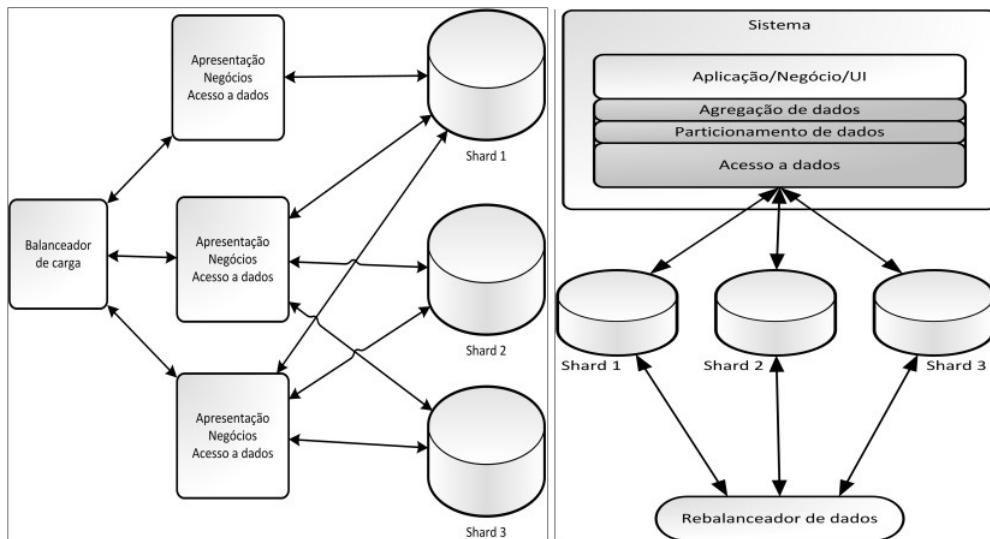


Figura 4. SNA com *Sharding* e a Arquitetura de um sistema com *Sharding*

Um fato relevante é que quando há a necessidade de escalar um banco de dados, ou a camada de dados de uma aplicação, o que geralmente se faz é adquirir computadores com maior capacidade de processamento, escalando verticalmente, e como se sabe esta abordagem é cara e limitada. Nesta situação, a opção é escalar o sistema horizontalmente, realizando uma distribuição dos dados.

Realizar *Sharding* de um banco de dados, como mostra a Figura 4, significa dividi-lo em instâncias menores, particionando e distribuindo os dados em um número de servidores de banco de dados. *Sharding* é um método de particionamento horizontal de dados que tem como objetivo a escalabilidade horizontal e o desempenho a um preço acessível. O rebalanceamento de dados nos *shards* pode ser implementado como um aplicativo separado, que monitora os *shards* e faz o rebalanceamento dos dados.

3.5. Arquitetura BASE

O padrão BASE (Figura 5) permite a construção de sistemas nos quais se troca a consistência de dados por escalabilidade ou disponibilidade, através da construção de um sistema que é basicamente disponível, lidando com dados ligeiramente

desatualizados e eventualmente consistentes. O uso de BASE em um sistema significa que o mesmo terá as seguintes propriedades:

- Basicamente disponível: o sistema terá tolerância a falhas parciais, mas não a falhas totais;
- Soft state: o sistema trabalhará com dados ligeiramente desatualizados;
- Consistência eventual: a propriedade de consistência, como especificada no Teorema CAP, não será respeitada, isto é, uma escrita em determinado dado não será visível a todo o sistema imediatamente, mas eventualmente o novo valor do dado será propagado para todo o sistema e todos os clientes e nós do sistema verão o mesmo dado [Vogels, 2008].

Talvez o uso mais conhecido do padrão BASE seja o sistema DNS (Domain Name System). Os dados são particionados entre vários servidores.

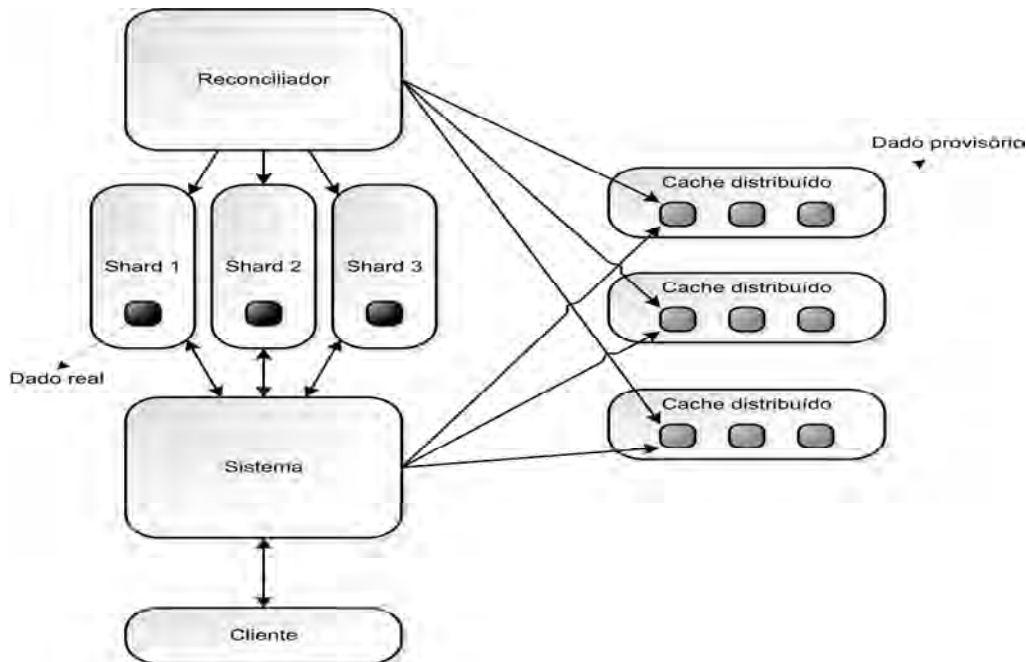


Figura 5. Estrutura de uma Arquitetura BASE com cache distribuído

4. Conclusão

Recentemente muita importância tem sido dada a sistemas escaláveis, pois praticamente todo sistema desenvolvido nos dias de hoje tem entre seus requisitos não funcionais a escalabilidade. O padrão Arquitetura *Shared Nothing* objetiva a construção de sistemas com escalabilidade linear através do uso de várias instâncias independentes do sistema evitando o aparecimento de gargalos. O padrão *Sharding* é uma alternativa para escalar os dados de um sistema, onde existem umas partes mais difíceis de serem escaladas horizontalmente. BASE é um padrão amplo que ataca vários problemas, onde se troca uma consistência forte dos dados por disponibilidade e tolerância a partições e se ganha em consequência escalabilidade e desempenho.

5. Referências

- Amdahl, G. M., (1967). Validity of the single processor approach to achieving large scale computing capabilities. In AFIPS '67 (Spring): Proc. of the April 18-20, spring joint computer conference, pp. 483–485, Atlantic City, New Jersey. ACM.
- Brewer, E. A. (2000). Towards robust distributed systems (abstract). In PODC. '00: Proc. of the nineteenth annual ACM symposium on Principles of distributed computing, p. 7, Portland, Oregon, United States. ACM.
- Demers, A., Petersen, K., Spreitzer, M., Terry, D., Theimer, M., e Welch, B., (1994), The Bayou Architecture: Support for Data Sharing among Mobile Users.
- Gamma, E., Helm, R., Johnson, R., e Vlissides, J. (1995). Design patterns: elements of reusable object-oriented software. Addison-Wesley Professional.
- Gray, C. e Cheriton, D., (1989). Leases: an efficient fault-tolerant mechanism for distributed file cache consistency. In SOSP '89: Proceedings of the twelfth ACM symposium on Operating systems principles, pp. 202–210, ACM.
- Porto, Ivens O. (2009), Dissertação de Mestrado sob o título “Padrões e diretrizes arquiteturais para escalabilidade de sistemas”, FACOM-UFU.
- Vogels, W. (2008). Eventually Consistent. Queue, 6(6):14–19.
- Williams, L. G. and C. U. S (2004). Web Application Scalability: A Model Based Approach. Technical report, Computer Measurement Group Conference (CMG'04).

Otimização de circuitos usando Simulated Annealing

Vinicius Veroneze R. Costa¹, Ivan S. Sendin¹

¹Departamento de Ciência da Computação – Universidade Federal de Goiás
Campus Catalão (UFG)
75.704-020 – Catalão – GO – Brazil

veroneze@gmail.com, ivansendin@yahoo.com

Abstract. *In this paper, the Simulated Annealing optimization method is studied in a introductory way, seeking to build an initial idea about its use in modeling for circuit design optimization. Graphics and the resulting configuration are shown proving the efficiency, flexibility and reliability of the method.*

Resumo. *O método Simulated Annealing é estudado de forma introdutória neste artigo, buscando construir uma ideia inicial sobre seu uso e desenvolvimento numa modelagem para otimização de design de circuitos. São mostrados gráficos e a configuração resultado comprovando a eficácia, flexibilidade e confiabilidade do método.*

1. Introdução

Na metalurgia, annealing (recozimento) é o processo de aquecer um metal e em seguida resfriá-lo lentamente para reduzir seus defeitos, otimizando sua estrutura. Aquecer o metal faz com que os átomos movimentem-se aleatoriamente, mas controlando o aquecimento e resfriamento, aumenta-se a probabilidade de que os átomos se reorganizem em cristais maiores, com menor energia interna e, conseqüentemente, com menos defeitos.

A versão para computadores do Simulated Annealing imita a versão metalúrgica, determinando menores níveis de energia interna para a função objetivo, muitas vezes chamada de função custo. O algoritmo inicia com uma temperatura elevada que lentamente é reduzida enquanto se realizam modificações elementares, buscando uma aproximação do mínimo global da função objetivo.

Existem vários problemas de otimização que se tornam intratáveis quando se usam métodos combinatórios, principalmente se a quantidade de elementos apresenta grande dimensão. Para estes problemas, uma metaheurística para otimização estocástica, como Simulated Annealing, é altamente efetiva.

2. O Método

A execução do algoritmo Simulated Annealing ocorre da seguinte maneira:

- Dada uma configuração inicial S_i
- Obtem-se a temperatura inicial T
- Enquanto o sistema não está solidificado, repita:
 - Modificação elementar em S_i obtendo S_j
 - Determinação da variação de energia ΔE

- Regra de aceitação de Metropolis:
 - Se $\Delta E \leq 0$, a modificação é aceita.
 - Se $\Delta E \geq 0$, a modificação pode ser aceita com probabilidade $e^{\frac{-\Delta E}{T}}$
- Redução paulatina de T
- Retorne a configuração otimizada.

De acordo com [Dreo 2006], a temperatura pode ser um simples elemento controlável do algoritmo ou pode-se criar uma função que a determine, baseada nos parâmetros do problema em questão. Na maioria dos casos, pode-se determinar a temperatura empiricamente sem perdas na qualidade de execução do algoritmo. Para a redução da temperatura, é recomendável que se aplique

$$T = \alpha * T, | 0,5 \leq \alpha \leq 0,9 \quad (1)$$

Para a modificação elementar em S_i , a variação de posição de um elemento aleatório pode ser suficiente, mas inverter a posição de dois elementos também pode ser uma boa escolha.

Para a determinação da variação de energia ΔE , primeiro define-se a função objetivo $f(s)$, que terá seu valor determinado com base em uma configuração passada como parâmetro. Assim, poderemos obter a variação de energia:

$$\Delta E = f(S_j) - f(S_i) \quad (2)$$

Quando $\Delta E \leq 0$, a modificação é aceita e faz-se $S_i = S_j$, mas quando $\Delta E \geq 0$, S_j só será aceito se $e^{\frac{-\Delta E}{T}} \geq R(0,1)$, onde $R(0,1)$ é um número real entre 0 e 1 gerado aleatoriamente. Note que quanto maior o valor de T , maior a probabilidade de S_j ser aceito.

3. O problema

Suponha um circuito com $n \times m$ elementos iguais e que cada elemento tenha v conexões com outros elementos. A função objetivo pode ser determinada como:

$$f = \sum_{i=0}^q \sum_{j=0}^v \sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} \quad (3)$$

onde q é a quantidade de elementos do circuito e v é a quantidade de “vizinhos” de cada circuito, ou seja, a quantidade de elementos conectados ao elemento q , (x_i, y_i) são as coordenadas do elemento q e (x_j, y_j) são as coordenadas do vizinho v atual. Note que a função custo é simplesmente a soma da distância euclidiana entre os elementos e seus elementos vizinhos, ou seja, a soma do comprimento das conexões entre os componentes do circuito.

Uma aplicação para o método Simulated Annealing é a otimização de circuitos eletrônicos, para obter uma melhor disposição dos elementos e diminuir o custo com o material utilizado nas conexões entre eles.

Foi criado um circuito de 5 x 5 elementos, e com 80 conexões, conforme a (Figura 1). Os quadrados vermelhos representam os componentes do circuito e as linhas azuis representam as conexões.

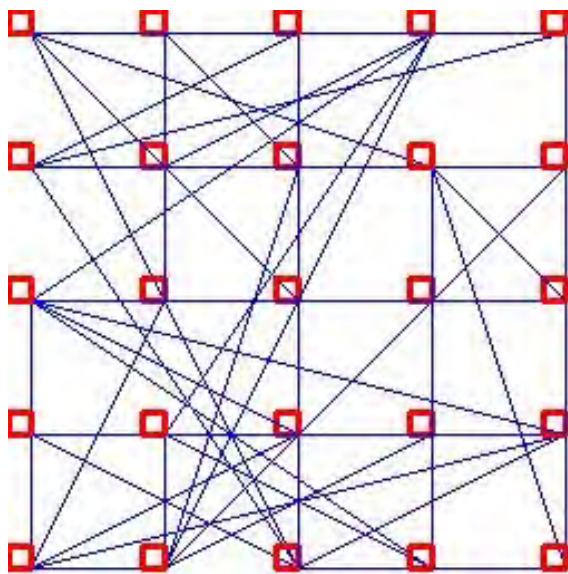


Figure 1. Circuito com a configuração inicial

Esta configuração inicial apresentou custo de 143,73 para as 80 conexões existentes, o que totaliza 1,797 de custo médio por conexão. Foi executado o algoritmo do Simulated Annealing com $T = 100,85$, $\alpha = 0,9$ e foi determinado que a configuração inicial estaria solidificada após 10.000 iterações. Em aproximadamente 18 segundos o algoritmo retornou uma configuração otimizada.

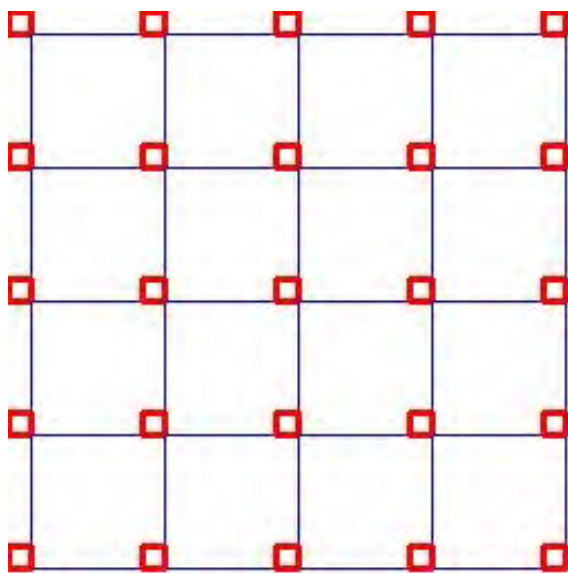


Figure 2. Circuito com a configuração otimizada

Na configuração otimizada, o custo total foi de 80, levando a 1 como custo médio. Perceba que esta configuração se trata de um mínimo global, e não somente um mínimo local ou uma aproximação ao mínimo global (Figura 2).

Um gráfico da execução é mostrado na (Figura 3).

É interessante observar que nas primeiras iterações ocorre um aumento na função

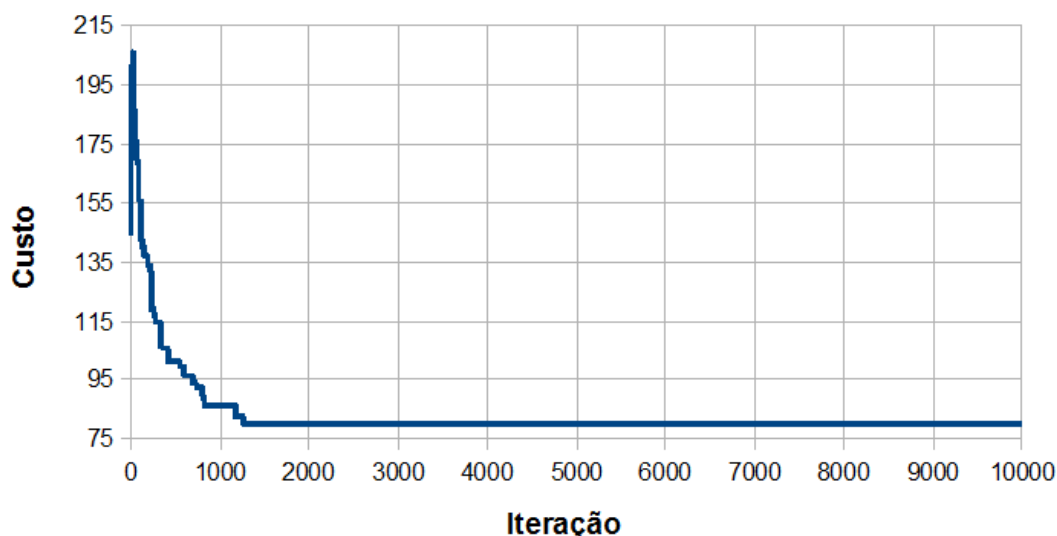


Figure 3. Gráfico Iteração x Custo

custo devido ao fato de os valores de T ainda serem altos. Esses valores aumentam a probabilidade de ser aceita uma solução que não possua custo menor que o atual, ou seja, $e^{\frac{-\Delta E}{T}} \geq R(0,1)$ será verdadeiro mais vezes para valores maiores de T . Detalhou-se o início da execução do algoritmo na (Figura 4).

Até a 10ª iteração o algoritmo seleciona uma configuração com custo mais alto que o inicial. Entre as iterações 10 e 40 são escolhidas configurações com custos menores, mas que são descartadas em detrimento a configurações com custo maior. Isso se trata da capacidade do algoritmo de fugir dos mínimos locais, na busca pela aproximação ao mínimo global.

Em circuitos de maior dimensão, o algoritmo retornou mais vezes uma configuração com custo aproximado ao mínimo global da função ao invés do mínimo global exato, como na configuração com 5x5 elementos. Em uma das execuções, foi gerado um circuito 10x10 (100 elementos, 360 conexões) que apresentou custo 1734. Após a aplicação do Simulated Annealing, foi retornada uma configuração com custo de 421, que representa 1,17 de custo por conexão. O mínimo global retonaria um custo por conexão com valor 1,0. A execução do algoritmo, nesse caso, se deu em 73 segundos.

4. A implementação

Apesar de não haver grandes complicações para a implementação do algoritmo, alguns desafios surgiram e relatam-se a seguir.

Foi escolhida a linguagem Python para a implementação, principalmente pela sua capacidade de manipulação de listas. Na (Figura 5) está o código da função que realiza o Simulated Annealing.

A função recebe 4 parâmetros:

- a é a configuração inicial do circuito

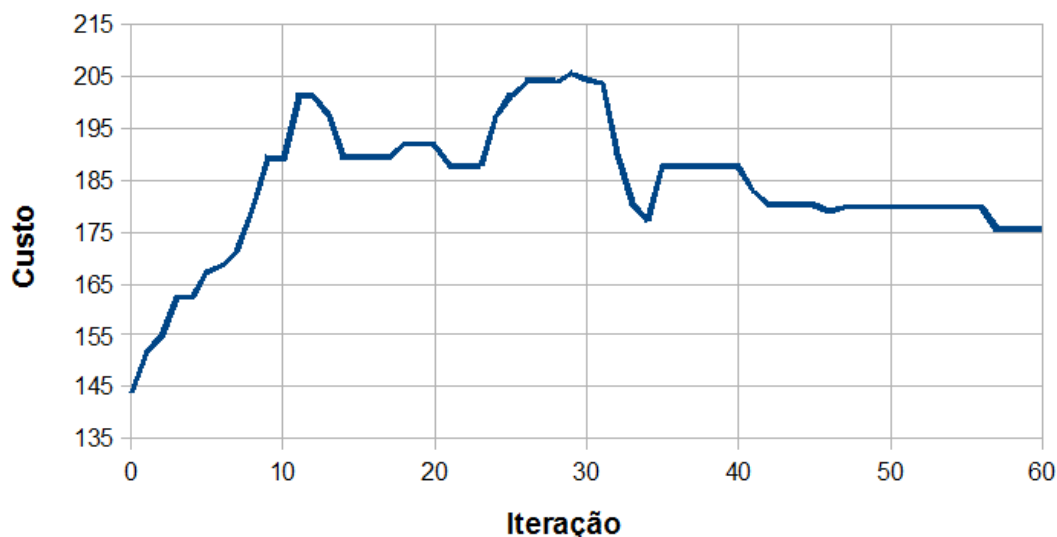


Figure 4. Primeiras 60 iterações

- t é a temperatura inicial
- n é a quantidade de iterações
- α é o fator de decremento da temperatura

O algoritmo repete por n iterações, calculando o delta, que é a diferença entre o custo da configuração modificada e o custo da melhor configuração atual. Caso delta seja menor que 0, significa que a configuração modificada é melhor que a atual, então temos uma nova melhor configuração. Caso delta não seja menor, se $e^{\frac{-\Delta E}{T}} \geq R(0, 1)$, a configuração é aceita mesmo tendo custo maior. Em seguida, t é decrementado proporcionalmente a α . Ao final, o algoritmo retorna a melhor configuração encontrada.

Note na linha 8 a chamada à função perturba. Essa função é a responsável por trocar a posição de 2 elementos aleatórios da configuração.

É importante ressaltar o uso constante de `copy.deepcopy()`. Como o Python trabalha com referências, principalmente com as listas, que são objetos, fazer $S_i = S_j$ está criando uma referência somente, ou seja, quando chama-se `perturba(S_j)`, na verdade S_i está sendo passado como argumento, o que não deixará outra configuração para comparar o custo, por isso a necessidade do uso de `copy.deepcopy()`. Apesar do alto custo computacional dessa função, o algoritmo foi executado rapidamente.

5. Conclusão

Neste trabalho buscou-se introduzir o algoritmo do Simulated Annealing, usando uma aplicação em otimização de design de circuitos. Foi observado que ele consegue encontrar o mínimo global da função custo na maioria das vezes, bastando ajustar os argumentos de entrada, principalmente a temperatura e a quantidade de iterações.

Nos próximos trabalhos pretende-se desenvolver o Simulated Annealing para refinamento de proteínas, com estrutura tridimensional e aplicações na Bioinformática. Um

```

1 def sa(a, t, n, alfa):
2
3     Si = copy.deepcopy(a)
4     Sj = copy.deepcopy(a)
5
6     for i in range(n):
7         ci = getCusto(Si)[0]
8         Sj = copy.deepcopy(perturba(copy.deepcopy(Si)))
9         cj = getCusto(Sj)[0]
10
11        delta = cj - ci
12
13        if delta < 0:
14            Si = copy.deepcopy(Sj)
15
16        else :
17            r = random.random()
18
19            if math.exp(-delta / t) > r:
20                Si = copy.deepcopy(Sj)
21
22        t = t * alfa
23
24    return Si

```

Figure 5. Código da função principal do Simulated Annealing

dos primeiros passos será o desenvolvimento de uma função para determinar a temperatura inicial do modelo para substituir a temperatura empírica que utilizou-se nesse trabalho.

References

- Dreo, J. (2006). *Metaheuristics for Hard Optimization*. Springer.
- Gielen, G., Walscharts, H., and Sansen, W. (1990). Analog circuit design optimization based on symbolic simulation and simulated annealing. *Solid-State Circuits, IEEE Journal of*, 25(3):707–713.
- Ji-Yang, Q. (2010). Application of improved simulated annealing algorithm in facility layout design. In *Control Conference (CCC), 2010 29th Chinese*, pages 5224–5227.
- Kumar, M. (2012). An information guided framework for simulated annealing. In *American Control Conference (ACC), 2012*, pages 827–832.

Considerações sobre um desenvolvimento acessível por meio do registro de *Design Rationale*

Humberto Lidio Antonelli¹, Thiago J. Bittar¹, Luanna L. Lobato¹

¹ Departamento de Ciências da Computação – Universidade Federal de Goiás (UFG)
Avenida Dr. Lamartine Pinto de Avelar, 1120 – Setor
Universitário – 75.704-020 – Catalão – GO – Brazil

humberto.antonelli@gmail.com, jabur@icmc.usp.br, lll@catalao.ufg.br

Abstract. *The Internet is an important means of transmitting information, providing interaction between people, in order to conduct business, education and government services. However, some accessibility problems, related to Web content, have been contributed to difficult the access by the users, particularly those that present some disability. In this sense, this paper discusses issues about accessibility, presenting a study on the system accessible development, using Design Rationale as way to capture the best project decisions, which can be reused as lessons learned about accessibility considerations, that need be considered during project development accessible.*

Resumo. *A Internet é um importante meio de transmissão de informação, proporcionando interação entre pessoas, afim de realizar negócios, educação e serviços governamentais. Porém, alguns problemas de acessibilidade, relacionados ao conteúdo Web, têm contribuído para dificultar o acesso pelos usuários, principalmente aqueles com algum tipo de deficiência. Neste sentido, o presente artigo aborda questões sobre acessibilidade, apresentando um estudo sobre o desenvolvimento de um sistema acessível, utilizando Design Rationale como forma de capturar as melhores decisões de projeto, que podem ser reutilizadas como lições aprendidas sobre as considerações de acessibilidade, que precisam ser consideradas durante o desenvolvimento de um projeto acessível.*

1. Introdução

Atualmente, devido a popularização da Internet, grande parte das tarefas diárias passaram a ser desempenhadas por meio da Web, tornando-a indispensável na rotina da maioria das pessoas. Desse modo, a Interface Humano-Computador (IHC), área responsável pelos aspectos que evoluem a relação homem-máquina, tornou-se essencial no desenvolvimento de qualquer aplicação Web, buscando alcançar sempre uma interface amigável, de fácil entendimento e, principalmente, acessível.

Para tanto, a acessibilidade é uma área que merece destaque, a qual possibilita que qualquer pessoa possa usufruir dos benefícios de uma vida em sociedade, ou seja, é a condição de participar de todas as atividades, até aquelas que incluem o uso de produtos, serviços e informação, com o mínimo de restrições possível [ABNT 2004]. Na Web, o uso de recursos também não deve ser diferente, para isso diversos órgãos criaram diretrizes e modelos que norteiam o processo de inclusão de acessibilidade de maneira padronizada e fácil, a fim de que a Web seja acessível para todos.

O processo de inclusão de acessibilidade em aplicações Web exige que, além da conscientização sobre a importância do tema entre os desenvolvedores e demais responsáveis por sua concepção, também seja realizada avaliações por meio de ferramentas e técnicas de acessibilidade. Neste sentido, é importante ressaltar que a busca por acessibilidade deve ser um processo contínuo, visto que a ausência de acessibilidade gera consequências, como, por exemplo, grupos de pessoas com dificuldades para acessar e obter informações [Pinto 2009].

Esta artigo apresenta considerações sobre o desenvolvimento de um sistema Web acessível, para utilização real pelo Departamento de Ciências da Computação da Universidade Federal de Goiás (UFG), Câmpus Catalão. Por meio desse desenvolvimento visa-se identificar problemas que possam dificultar ou impedir o acesso aos recursos disponíveis no sistema para qualquer pessoa com deficiência e implementar soluções que resolvam tais problemas, fazendo considerações sobre as barreiras de acesso identificadas e como foram reduzidas ou não.

O grande propósito deste trabalho é retratar a importância de um desenvolvimento Web acessível, baseado nas diretrizes de acessibilidade, mostrando considerações aplicadas sobre o desenvolvimento de um sistema real. Neste sentido, busca-se com este trabalho possibilitar a contribuição para o incentivo do processo de acessibilidade em sistemas Web, uma vez que a possibilidade de acesso é direito de vários tipos e perfis de usuários e, assim, deve ser garantido a todos.

Diante do exposto, a relevância da pesquisa consiste em estimular os desenvolvedores a adaptarem seus projetos de interface para pessoas com deficiência, aproximando os sistemas ao cenário ideal de uso, uma vez que considera as questões de acessibilidade. Assim, neste trabalho são destacadas as premissas em relação às lições aprendidas por meio do desenvolvimento de um sistema para uso real.

Este artigo está organizado da seguinte forma: a Seção 2 descreve sobre os conceitos de acessibilidade empregados no desenvolvimento deste trabalho, na Seção 3 são apresentados os conceitos de *Design Rationale*. A metodologia utilizada é abordada na Seção 4. A Seção 5 apresenta uma breve descrição da ferramenta desenvolvida a partir das diretrizes de acessibilidade. Os resultados e considerações sobre o desenvolvimento são apresentados na Seção 6. Por fim a Seção 7 apresenta as conclusões.

2. Acessibilidade na Web

A acessibilidade na Web visa garantir que pessoas com deficiência possam perceber, entender, navegar e interagir com a Web, para que, então, livre de barreiras, elas também possam contribuir [Rutter et al. 2006]. Acessibilidade na Web também beneficia outras pessoas, incluindo pessoas idosas com capacidades de movimentação reduzida devido ao envelhecimento. Além disso, engloba todas as deficiências que afetam o acesso à Web, seja visual, auditivo, da fala, físicas, cognitivas e/ou neurológicas [W3C 2005].

Embora o acesso as aplicações Web pelas pessoas com deficiência seja o principal foco da acessibilidade na Web, ela também beneficia pessoas sem deficiência [Rutter et al. 2006]. Por exemplo, a criação de sites que sejam flexíveis para atender as necessidades de diferentes usuários aumenta a usabilidade em geral e, permite que as pessoas sem deficiência usem esses sites de acordo com suas preferências, como, por exemplo, o uso de qualquer navegador.

Com o objetivo de tornar a Web acessível a um número cada vez maior de pessoas e de levá-la ao potencial máximo de interoperabilidade, o W3C criou a WAI, que mantém grupos de trabalho elaborando conjuntos de diretrizes e recursos que ajudam garantir a acessibilidade do conteúdo Web às pessoas com deficiências, ou que acessam a Web em condições especiais de ambiente, equipamento, navegador e outras ferramentas Web [W3C 2005].

As diretrizes propostas pelo W3C compõe um conjunto de três documentos como recomendações, as quais descrevem os recursos necessários para alcançar diferentes níveis de acessibilidade, além de incluir listas de referência e técnicas de implementação, os quais são:

- *Web Content Accessibility Guidelines (WCAG)*: é voltada para o conteúdo das páginas Web, visto que explica como tornar o conteúdo de sites mais acessível às pessoas com deficiência.
 - WCAG 1.0: é composta de recomendações gerais e pontos de verificação específicos que são divididos em 3 níveis de prioridade.
 - WCAG 2.0: atualização da WCAG 1.0 para abranger tecnologias mais avançadas da Web.
- *Authoring Tool Accessibility Guidelines (ATAG)*: é destinada para fabricantes de ferramentas de criação de conteúdo, pois fornece orientações para que possam implementar funcionalidades em seu produto que ajude ao desenvolvedor de conteúdo a obedecer as recomendações descritas na WCAG.
- *User Agent Accessibility Guidelines (UAAG)*: é voltada para desenvolvedores de navegadores e reprodutores multimídia, incluindo também alguns aspectos das tecnologias assistivas. Explica como desenvolver navegadores acessíveis para pessoas com deficiência.

Para o desenvolvimento deste trabalho foram utilizadas as diretrizes propostas pela WCAG 2.0, visto que é o documento mais atual que norteia o desenvolvimento de aplicações Web acessíveis [W3C 2008].

Enfim, para que a acessibilidade na Web seja de fato alcançada é importante que os desenvolvedores de conteúdo ou sistemas Web, assim como seus mantenedores, estejam atentos às recomendações de acessibilidade proposta por documentos nacionais e internacionais, em especial as diretrizes do W3C, como também às diferentes características e limitações dos usuários [Pupo et al. 2006, Nicácio 2010].

3. Design Rationale

Neste trabalho, o conceito de *Design Rationale* (DR) foi selecionado, para auxiliar no registro da tomada de decisões em relação ao desenvolvimento do sistema, principalmente, no que se diz respeito ao emprego dos conceitos de acessibilidade na Web, visto que se pretende construir uma base útil para a gestão do conhecimento sobre o projeto, bem como garantir o gerenciamento das decisões tomadas no decorrer do seu desenvolvimento.

Na literatura encontram-se várias definições envolvendo DR, sendo algumas delas apresentadas a seguir:

Segundo Lee (1997) e Lara (2005), os DRs são ferramentas importantes, pois incluem não somente as razões por trás de uma discussão de design, mas também as

justificativas para as mesmas, outras alternativas que foram consideradas pela equipe, as vantagens e desvantagens que foram avaliadas, e principalmente as argumentações que levaram às tomadas de decisão.

Para MacLean *et al.* (1989) o DR é o conjunto de opções que são selecionadas para o projeto final que descrevem o artefato, em que as alternativas e razões para as escolhas fornecem argumentos que apoiam e ajudam a compreensão das escolhas realizadas.

Burge e Brown (2000) caracterizam o DR como inestimável, pois colabora com a revisão, manutenção, documentação, avaliação e aprendizagem do projeto, já que o mesmo armazena as decisões tomadas, bem como os motivos que as definiram. Para Faria e Bittar (2011) o DR ajuda também na identificação de premissas inadequadas, auxiliando os projetistas a perceber alternativas possíveis sobre decisões importantes.

Enfim, o DR contém o registro de todas as informações acerca do projeto, o que configura uma base de conhecimentos que pode ser utilizada para futuras consultas. O uso do DR melhora a compreensão do desenvolvimento de um software, o que facilita sua manutenção, sua evolução e o seu reuso em novos projetos. Além disso, proporciona uma melhor comunicação entre os membros de uma equipe de desenvolvimento, fazendo com que durante o processo de software seja possível entender quais foram as decisões críticas tomadas, quais alternativas foram investigadas, entre outras informações que venham a ser necessárias para o melhor entendimento do projeto, evitando que enganos posteriores sejam cometidos.

O registro e uso do DR agrega vários benefícios ao processo desenvolvimento de software, dentre os quais pode-se citar:

- auxilia na resolução de problemas, pois o DR fornece informações sobre o raciocínio por trás do projeto [MacLean et al. 1989].
- facilita a manutenção e evolução do artefato, visto que permite aos *stakeholders* um entendimento mais claro e completo acerca do projeto [MacLean et al. 1989];
- colabora com a comunicação entre os *stakeholders*;
- facilita o aprendizado do projeto, contribuindo para que novos *stakeholders* possam facilmente ser integrados na equipe [Conklin and Burgess-Yakemovic 1996];

Em suma, segundo Santos (2007), o objetivo maior do registro e uso do DR é a produção de artefatos mais completos, em menor tempo, com menos esforço, mais fáceis de serem mantidos e evoluídos e com melhor qualidade.

4. Metodologia

A metodologia empregada nesse trabalho levou em consideração aspectos relativos ao aprendizado, visando apresentar o universo da pesquisa teórica e aplicada, envolvendo as bases da utilização do DR como uma ferramenta de auxílio no processo de tomada de decisões no decorrer da construção, manutenção e evolução de um sistema Web, proporcionando o levantamento de consideração sobre o emprego dos conceitos de acessibilidades propostos pelas diretrizes do W3C.

Para o levantamento de considerações sobre um desenvolvimento acessível por meio do registro de DR, optou-se pelo desenvolvimento de um sistema real. A definição de qual sistema construir partiu do um levantamento de problemas enfrentados pelos eventos acadêmicos (congressos, simpósios, encontros, etc) que aconteceram no âmbito da

UFG, Câmpus Catalão, no qual se destacou a falta de uma ferramenta que pudesse estar gerenciando as informações geradas com a realização do evento.

Diante do exposto, foi proposta a construção de um sistema Web genérico, para que qualquer tipo de evento acadêmico possa encontrar a melhor estrutura online para efetuar seu gerenciamento, com uma plataforma inteligente e simplificada para favorecer a organização do evento, possibilitando um modelo fácil de administração.

O desenvolvimento deste tipo de sistema envolve diversos fatores, como a elaboração de formulários, *grid* de dados, disponibilização de calendário, tabelas, relatórios, entre outros. Desse modo, torna-se um meio propício para extrair considerações práticas sobre o processo de inclusão de acessibilidade no sistema.

A construção deste sistema está sendo baseada na utilização da metodologia de desenvolvimento “beta perpétuo”, no qual o sistema ficará sempre em desenvolvimento, evoluindo com a colaboração dos próprios usuários. O DR está sendo utilizado para armazenar e gerenciar a tomada de decisões, incluindo as decisões de acessibilidade, que ocorre durante todo o processo de desenvolvimento, além de colaborar com a documentação e manutenção do projeto.

O registro do DR conta com o apoio da ferramenta AccessibilityUtil¹, que é um repositório colaborativo no qual possibilita uma melhor interação entre os membros da equipe, bem como o armazenamento de todas as opiniões acerca de determinada decisão. Além disso, a ferramenta provê sugestões de acessibilidade, baseadas no uso de diretrizes da WCAG 2.0, a serem adotadas em cada tópico, as quais são filtradas pelo uso de palavras-chaves. Essas sugestões vem colaborando para que a equipe consiga tornar menos dispendioso e oneroso o processo de tomadas de decisão para confecção de aplicativos acessíveis por fazer uso dessas diretrizes, juntamente com o reuso das informações já capturadas e avaliadas anteriormente.

5. A ferramenta SIADE

Por se tratar de um trabalho com fins acadêmicos elaborado dentro de uma instituição pública de ensino, o desenvolvimento desta ferramenta foi baseada no uso de tecnologias denominadas *open source*. Assim, a ferramenta SIADE foi desenvolvida utilizando a linguagem PHP, para a programação para a aplicação no lado do servidor, e com o auxílio do *framework* Javascript Dojo Toolkit, para a manipulação de eventos na interface. Um ponto importante para escolha deste *framework* foi por ele ser projetado para ser acessível, desde a sua concepção.

O Sistema Gerenciador de Banco de Dados (SGBD) escolhido para esta ferramenta foi o MySQL. Entretanto, ressalta-se que a aplicação tem suporte a diferentes banco de dados. Outra característica da ferramenta é ser multiplataforma, visto que as tecnologias utilizadas são preparadas para funcionarem em qualquer sistema operacional.

A ferramenta está disposta em módulos, agrupando um conjunto de funcionalidades disponíveis. Uma das vantagens dessa modularização é a possibilidade oferecida ao administrador do sistema de apontar as permissões de acessos dos usuários para cada módulo. De acordo com as funcionalidades idealizadas para compor a ferramenta, foi

¹<http://www.accessibilityutil.com/>

definido o modelo relacional da base de dados por meio da criação de 37 tabelas com relacionamentos entre si. A disposição da base de dados permite uma melhor adaptabilidade, expansão e evolução do sistema.

6. Resultados e considerações

A ferramenta SIADÉ está sendo utilizada para administração do X Encontro Anual de Computação (EnAComp)². O volume de informações geradas nas duas últimas edições deste evento, também foram migradas com sucesso para a ferramenta, possibilitando que os participantes das edições anteriores pudessem ter acesso às suas informações, principalmente, a emissão de seus certificados. Atualmente, a ferramenta possui ao todo mais de 630 inscrições efetuadas, sendo 150 destas referentes ao X EnAComp.

Como resultados principais, destacou-se a utilização com sucesso e satisfatória da ferramenta *AccessibilityUtil* pela equipe no gerenciamento das decisões tomadas durante o desenvolvimento do sistema, com ênfase na gestão de acessibilidade. A captura do DR em todas as fases deste projeto tem colaborado com a tomada de decisões proporcionando um produto final adequado às diretrizes do WCAG 2.0, bem como contribuído para uma comunicação efetiva entre os integrantes da equipe.

Com o uso da ferramenta *AccessibilityUtil* é possível gerenciar as decisões em forma de tópicos contendo o registro das opiniões dos integrantes da equipe, em forma de comentários, e o status de discussão em que se encontra, ou seja, se a equipe chegou a um acordo e concluiu determinada questão ou ainda está em discussão [Bittar et al. 2011].

As considerações sobre um desenvolvimento acessível podem ser observadas nos comentários de cada tópico, ou seja, por meio dos comentários que é exposta toda argumentação necessária para o correto emprego de acessibilidade no projeto³.

A Tabela 1 apresenta a relação de alguns tópicos criados no *AccessibilityUtil* para gerenciamento da tomada de decisões, com ênfase no aspecto de acessibilidade.

Tabela 1. Relação de alguns tópicos cadastrados

Título	Comentários	Status
Linguagem de programação a ser utilizada	4	Concluído
Perfil de usuários - público alvo	4	Concluído
Linguagem narrativa a ser utilizada	3	Concluído
Ícones, quais usar?	4	Concluído
Processo de desenvolvimento e acessibilidade	2	Em discussão
Quais testes serão feitos e em que momento?	3	Em discussão
Versão de HTML a ser utilizada	2	Em discussão
Cores, quais usar?	4	Concluído
Frameworks de javascript	3	Concluído
Como será feita a identificação de requisitos?	2	Em discussão
Validação de CSS	3	Em discussão

Apesar do sistema já estar em uso, ainda passa por constantes alterações, graças à

²<http://www.enacomp.com.br/siade/>

³Todas essas informações podem ser acessadas por meio do link: <http://www.accessibilityutil.com>

utilização da metodologia beta perpétuo, no qual o sistema evolui continuamente a partir da colaboração dos usuários que são os principais testadores [Bianchini 2008]. A ideia é engajar os usuários como testadores em tempo real para que se possa saber qual e como tal serviço está sendo usado, identificando problemas que estejam ocorrendo para que correções possam ser feitas, como também a criação de novas funcionalidades e melhorias de funcionalidades já existentes. Tais informações geradas são armazenadas e utilizadas na tomada de decisões.

No que diz respeito a ocorrência de reclamação ou observação pelos usuários sobre a falta de acessibilidade na ferramenta desenvolvida, não foi reportada nenhuma informação até o momento. Contudo, isto pode ser em decorrência do público alvo do evento no qual a ferramenta está sendo utilizada, ou seja, trata-se de um evento na área de computação, logo, presume-se que os participantes tenham mais facilidade em fazer uso da tecnologia. Neste caso, este fato é encarado como uma limitação dos resultados para este artigo.

7. Conclusão

A captura do DR durante todas as fases deste projeto colaborou com a tomada de decisões e o correto emprego de características de acessibilidade no desenvolvimento do sistema, bem como contribuiu para uma comunicação efetiva entre os integrantes da equipe.

O armazenamento das decisões em um repositório colaborativo facilitou o gerenciamento das informações no decorrer do desenvolvimento, além de possibilitar a reutilização das decisões em outros projetos ou na revisão de problemas semelhantes encontrados no projeto atual.

A inclusão de acessibilidade no desenvolvimento proposto por este trabalho não foi uma tarefa fácil, porém o aprendizado adquirido por meio do DR juntamente com o conhecimento aprofundado, sobre os conceitos de acessibilidade na Web, por alguns integrantes da equipe proporcionaram uma maior facilidade de inclusão do que em outros projetos. As considerações geradas com este trabalho podem servir de exemplo para outros desenvolvedores que desejam adaptá-las à seus projetos, a fim de alcançar também um desenvolvimento acessível.

A aplicação da metodologia beta perpétuo possibilita uma melhor adaptabilidade e aceitabilidade do sistema, visto que conta com a colaboração dos próprios usuários para que melhorias sejam feitas de acordo com as sugestões propostas pelos mesmos.

Referências

- ABNT (2004). *NBR 9050: Acessibilidade a edificações, mobiliário, espaços e equipamentos urbanos*. Associação Brasileira de Normas Técnicas, Rio de Janeiro.
- Bianchini, S. L. (2008). Avaliação de métodos de desenvolvimento de aplicações web. Mestrado em ciências de computação e matemática computacional, Instituto de Ciências Matemáticas e de Computação de São Carlos (ICMC/USP), Universidade de São Paulo, São Carlos - SP.
- Bittar, T. J., do Amaral, L. A., and Fortes, R. P. d. M. (2011). Accessibilityutil: a tool for sharing experiences about accessibility of web artifacts. In *Proceedings of the 29th*

- ACM international conference on Design of communication*, SIGDOC '11, pages 17–24, New York, NY, USA. ACM.
- Burge, J. and Brown, D. C. (2000). Reasoning with design rationale. In *Artificial Intelligence in Design '00*, pages 611–629, Netherlands. Kluwer Academic Publishers.
- Conklin, E. J. and Burgess-Yakemovic, K. (1996). Design rationale. chapter A process-oriented approach to design rationale, pages 393–427. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.
- de Lara, S. M. A. (2005). Um suporte à captura informal de design rationale. Mestrado em ciências de computação e matemática computacional, Instituto de Ciências Matemáticas e de Computação de São Carlos (ICMC/USP), Universidade de São Paulo, São Carlos - SP.
- Faria, F. B. and Bittar, T. J. (2011). Uso de design rationale para gestão de conhecimento em acessibilidade na web. In *Anais eletrônicos...*, Catalão-GO. Encontro Anual de Computação, UFG.
- Lee, J. (1997). Design rationale systems: understanding the issues. *IEEE Expert*, 12(3):78–85.
- MacLean, A., Young, R. M., and Moran, T. P. (1989). Design rationale: the argument behind the artifact. In *Proceedings of the SIGCHI conference on Human factors in computing systems: Wings for the mind*, CHI '89, pages 247–252, New York, NY, USA. ACM.
- Nicácio, J. M. (2010). *Técnicas de acessibilidade: criando uma Web para todos*. Edufal - Editora da Universidade Federal de Alagoas, Maceió - AL.
- Pinto, C. S. (2009). Avaliação da acessibilidade de sistemas web de comunicação e de gestão de grupos visando pessoas com deficiência visual total. Mestrado em informática, Departamento de Informática Aplicada, Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Rio de Janeiro - RJ.
- Pupo, D. T., Melo, A. M., and Ferrés, S. P. (2006). *Acessibilidade: discurso e prática no cotidiano das bibliotecas*. UNICAMP, Campinas - SP.
- Rutter, R., Lauke, P., Waddell, C., Thatcher, J., Henry, S., Lawson, B., Kirkpatrick, A., Heilmann, C., Burks, M., Regan, B., et al. (2006). *Web Accessibility: Web Standards and Regulatory Compliance*. Apresspod Series. Apress.
- Santos, D. R. G. (2007). Suporte ao registro e uso de decisões de projetos de aplicações para a web. Mestrado em informática, Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro - RJ.
- W3C (2005). Introduction to web accessibility. Disponível em: <http://www.w3.org/WAI/intro/accessibility.php>. Acesso em: 27 set. 2012.
- W3C (2008). Web content accessibility guidelines 2.0. Disponível em: <http://www.w3.org/TR/WCAG20/>. Acesso em: 03 set. 2012.

Autenticação Biométrica em Redes de Telecomunicação Aeronáutica

Rafael de Amorim Silva^{1,2}, Angelo Zanini¹

¹Departamento de Eletrônica - Instituto Tecnológico de Aeronáutica (ITA)
– São Jose dos Campos – SP – Brasil

{amorim, azanini}@ita.br

²Instituto de Computação
Universidade Federal de Alagoas (UFAL) – Maceió, AL – Brasil

rafael@ic.ufal.br

Abstract. *This paper proposes a network architecture that supports biometric authentication on aeronautical networks to provide more confident services related to air passenger communication. The chosen authentication process is the single logon. This architecture is based on overlay protocol between application and transport layers, obtaining agents that negotiate the authentication process. As a result, the architecture enables the development of various applications to assure information trafficked by aeronautical networks.*

Resumo. *Este artigo propõe uma arquitetura de rede que auxilie a autenticação biométrica em redes aeronáuticas para prover serviços mais confiáveis relacionados a comunicação para passageiros aéreos. O processo de autenticação adotado é o single logon. A arquitetura é baseada em um protocolo overlay entre as camadas de aplicação e transporte, possuindo agentes que negociam o processo de autenticação. Como resultado, a arquitetura permite o desenvolvimento de diversas aplicações para assegurar a informação trafegada em redes aeronáuticas.*

1. Introdução

Uma rede de telecomunicação aeronáutica é capaz de se comunicar com uma aeronave através de vários tipos de enlace de dados (e.g. satélites, comunicações aéreo-aéreo e aéreo-terrestres) [1]. Tais redes são projetadas para trabalhar com aplicações que exijam um alto nível ou baixo nível de segurança. A rede ATN oferece quatro tipos de serviços [2]: (i) sistema de tráfego aéreo (ATS); (ii) Controle operacional aéreo(AOC); (iii) Controle administrativo das linhas aéreas (AAC); e (iv) comunicação de passageiros aéreos (APC). Os serviços ATS e AOC exigem um forte nível de segurança, usados geralmente em comunicações operacionais. Já os serviços APC e AAC exigem baixos requisitos de segurança, usados geralmente em comunicações não operacionais.

Inicialmente, a indústria tem adotado o modelo OSI para comunicações aeronáuticas. Entretanto, iniciativas como o projeto NEWSKY [3] propõe a migração do modelo OSI para o modelo TCP/IP, ofertando uma nova geração de tecnologias que serão empregadas para uso aeronáutico. Embora hajam esforços na direção de uma rede de telecomunicação aeronáutica baseada na pilha atual da Internet, tal modernização ainda se

encontra em estágio inicial. Diversas aplicações serão desenvolvidas devido a aquisição deste novo padrão. Alguns trabalhos [4, 5] têm proposto alternativas para aumentar a segurança à nível de arquitetura para estas redes. Todavia, existem poucos trabalhos que lidem com aspectos relacionados a segurança das aplicações neste novo cenário aeronáutico. A arquitetura TCP não oferece a segurança adequada para o processo de autenticação, sendo que problemas como erros em canal, excessivos *handoffs*, falta de conectividade podem dificultar o uso de algoritmos de segurança. Para resolver tal problema, uma aplicação que captura uma credencial biométrica e a distribui para as estações aéreo-terrestres programadas em um plano de vôo estabelece comodidade para o usuário e aumenta o desempenho na rede.

Portanto, este artigo propõe uma arquitetura de rede projetada para suportar a autenticação biométrica em redes de telecomunicação aeronáutica que implementem serviços APC. Tal arquitetura oferece um protocolo *overlay* entre as camadas de aplicação e transporte, com agentes responsáveis na realização do processo de autenticação. Dentre os benefícios esperados, tal arquitetura propicia o desenvolvimento de inúmeras aplicações para a segurança de dados dos usuários que viajam em uma aeronave.

O trabalho é organizado da seguinte maneira: Na Seção 2 são apresentados alguns trabalhos relacionados. A Seção 3 apresenta as principais técnicas da autenticação biométrica. A Seção 4 descreve a arquitetura proposta neste trabalho, descrevendo o cenário, a pilha de protocolos e a arquitetura de agentes envolvida. A Seção 5 apresenta um estudo de caso para avaliar a arquitetura, definindo alguns exemplos de aplicações que se beneficiam da implantação desta arquitetura. A Seção 6 apresenta as considerações finais deste trabalho.

2. Trabalhos Relacionados

O projeto NEWSKY [3] foi pioneiro no desenvolvimento e análise da nova geração de tecnologias baseadas no modelo ATN/IPS que serão empregadas no cenário aeronáutico nos próximos anos. O escopo desta investigação inclui a rede aéreo-terrestre, considerando apenas os serviços que exijam um alto nível de segurança. Além do projeto NEWSKY, outros autores em [10] também propuseram arquiteturas de rede aeronáutica baseados no padrão ATN/IPS para comunicações aéreo-terrestres. Estes autores discutem problemas relacionados a mobilidade, segurança e integração em arquiteturas ATN. Autores em [10] investigam a aplicação da tecnologia Mobile IP em programas aeronáuticos e da agência espacial norte-americana, discutindo as características, capacidades e possíveis cenários de uso. Em particular, estes autores investigam a conectividade IPv4 e IPv6 entre *hosts* fixos e móveis através de *testbeds*. As questões relativas a segurança destas redes são discutidas em [4, 5].

3. Autenticação Biométrica

Identificação pessoal é o processo de associar um indivíduo particular a uma identidade. Tal identificação pode ser tanto na forma de verificação (autenticação), o qual autentica uma identidade exigida, ou reconhecimento (identificação), que determina a identidade de uma pessoa através da consulta de um banco de dados de pessoas conhecidas no sistema. Para identificar automaticamente uma pessoa, usa-se algo que ela possua em mãos (abordagem baseada em tokens), ou usa-se algo que ela saiba (abordagem baseada em conhecimento), ou usa-se alguma característica fisiológica ou comportamental (abordagem

baseada em biometria). A abordagem biométrica tem se destacado devido a sua capacidade de identificar, com maior probabilidade, a pessoa correta para autorização de acesso e recursos em um determinado sistema. As outras abordagens possuem uma capacidade inferior de identificar corretamente um usuário, pois os tokens podem ser perdidos, roubados, esquecidos, ou deixados em algum lugar. Além disso, um PIN pode ser esquecido ou pego por um impostor, falhando em distinguir entre uma pessoa autorizada e um impostor.

O processo de identificação em um sistema biométrico é dividido em duas etapas: (i) captura de uma característica biométrica através de um sensor e armazenada em um banco de dados para futuras comparações (conhecida como etapa de cadastramento das credenciais biométricas); (ii) processo de comparação em tempo real de ambas as credências (etapa de comparação). Na fase de cadastramento, a característica biométrica é escaneada por um sensor biométrico para adquirir a representação digital da característica. Depois, a representação digital é processada por um extrator de características para gerar uma representação expressiva e compacta (*template*). Este *template* pode ser armazenado ou em um banco de dados central ou em um cartão magnético (local). Na fase de reconhecimento, o leitor biométrico captura a característica do indivíduo e converte-o a um formato digital. Depois, processa-se essa captura pelo extrator de características para produzir a mesma representação como o *template*. Por último, a representação resultante é passada ao extrator de característica que compara-o novamente com o *template* para estabelecer a identidade do indivíduo [6].

Para a escolha adequada de uma determinada característica biométrica, deve-se levar em conta os seguintes fatores: universalidade, distinção, permanência, coletabilidade, desempenho, aceitabilidade e circumvenção [7].

3.1. Métodos de Autenticação Biométrica

Existem vários métodos para a captura de uma característica biométrica. Tais métodos se diferenciam através dos fatores de escolha citados na Seção 3. Estes métodos são: Reconhecimento de Face, Termograma Facial, Geometria das Mãos, Reconhecimento por Íris, Assinatura Humana, Fala Humana, Reconhecimento por Retina, Veias da Palma da Mão e Impressão Digital [7, 8].

A identificação biométrica por reconhecimento da face humana é uma das mais comuns características utilizadas para identificação pessoal. Tal identificação é baseada na localização e na forma de atributos como olhos, nariz, lábios, sobrancelhas, queixo e seus relacionamentos espaciais. Tal identificação é adotada por aplicações que funcionam em ambientes controlados ou fundos com imagem distorcida (e.g. identificação de presos e criminosos). Este modo de identificação requer iluminação especial, sendo de difícil reconhecimento caso a captura da imagem não tenha sido obtida pelo ângulo original da imagem.

O sistema de vasos sanguíneos existentes no lado interno da face permite a identificação de um indivíduo através de assinaturas faciais, devido a passagem de calor através do tecido facial. Por isso, um termograma facial identifica assinaturas faciais através dessa passagem, utilizando câmeras infra-vermelhas que capturam minúcias deste sistema. Tal método de identificação identifica unicamente uma característica do indivíduo e não é vulnerável a quaisquer tipos de disfarces. A estrutura vascular é mais rica em informação, fornecendo minúcias mesmo em ambientes com pouca iluminação.

O problema principal desta técnica é que é suscetível a mudanças de temperatura, estado emocional do indivíduo e ângulo da imagem obtida.

A técnica de identificação biométrica pelo uso da geometria das mãos humanas consiste em considerar a forma da mão, alturas e tamanhos dos dedos como característica biométrica. A técnica é simples, de fácil uso e bom custo-benefício. Fatores como umidade e pele úmida não influenciam na eficácia desta técnica. Todavia, a mesma possui baixa capacidade discriminativa. Outro problema é que a geometria da mão muda ao longo da vida, oferecendo uma baixa permanência da característica biométrica. Outros fatores como o uso de jóias e artrites podem dificultar a identificação.

A técnica de identificação biométrica pelo uso da íris é uma das técnicas mais eficazes, devido a sua longa permanência (a textura visual da íris estabiliza durante os dois primeiros anos de vida) e sua unicidade. Além disso, a característica biométrica é mais fácil de capturar do que a técnica de reconhecimento por retina. Dentre as suas vantagens, destacam-se a sua segurança, visto que é difícil cirurgicamente danificar a sua textura e a facilidade em detectar íris superficiais (fraudulentas).

A assinatura é o método mais tradicional de reconhecer a autenticidade de um indivíduo. Cada indivíduo possui um estilo único de assinatura, sendo que duas assinaturas deste mesmo indivíduo são geralmente idênticas. As variações dependem do seu estado físico e emocional. Existem 2 modos de identificação: Estática e dinâmica. O primeiro utiliza apenas características geométricas de uma assinatura. O segundo adota as características do primeiro e acrescenta fatores como aceleração, velocidade, pressão e trajetórias da assinatura.

A fala humana pode ser utilizada para identificação biométrica. O som característico de cada indivíduo é formado por diferenças na fala devido a fatores como forma ou tamanho invariante do trato vocal, boca, cavidades nasais e lábios. A verificação pode ser feita pelo uso de um texto (o indivíduo deve ler o texto e a comparação entre o texto falado e o armazenado em uma base de dados será realizado) ou independente de texto (o que oferece maior proteção contra fraudes). A aceitabilidade por parte dos usuários é excelente. Fatores como ruídos, estado físico e emocional do indivíduo pode influenciar em um baixo desempenho da técnica. Geralmente, tal técnica é adotada em aplicações com baixos requisitos de segurança devido a alta variabilidade existente na voz.

A técnica de identificação biométrica pelo reconhecimento retinal se caracteriza pelo uso de informações capturadas das veias que se situam abaixo da superfície retinal em um olho. O dispositivo de captura biométrica projeta um feixe com baixa intensidade visual em um olho para capturar uma imagem da retina. Portanto, o indivíduo deve se aproximar do dispositivo e direcionar a sua visão para um determinado ponto no campo visual. O problema desta técnica é que é intrusiva e de alto custo. Dentre as suas vantagens, a técnica oferece uma característica única em cada indivíduo e é estável.

A técnica de identificação biométrica pelo padrão das veias de uma palma da mão consiste na identificação de minúcias devido a absorção da luz em determinados comprimentos de onda que uma hemoglobina reduzida possui em relação a hemoglobina oxigenada [8]. Ou seja, emite-se uma luz em direção a uma palma da mão, sendo que uma luz com um determinado comprimento de onda será refletido. Uma vez que há absorção de luz pela hemoglobina reduzida, a luz refletida será identificada por uma câmera de

infravermelho, formando a imagem das veias da palma da mão. Dentre as suas vantagens, o método se apresenta como não invasivo, higiênico e único. Além disso, é uma técnica de difícil fraude e possui capacidade de detecção de vida sem precisar de dispositivos adicionais.

A técnica de identificação biométrica por impressão digital é a mais popular dentre as técnicas. O método consiste em capturar minúcias do padrão de cristas e sulcos na superfície de um dedo. As minúcias detectadas são transformadas em um vetor de características, sendo uma lista de localizações e outros atributos (orientação da curva, por exemplo). Posteriormente, compara-se a identificação biométrica com *templates* armazenados em uma base de dados. A técnica possui baixo custo de implementação. Dentre as desvantagens, as impressões digitais de uma parte da população podem ser inadequadas devido a questões genéticas, envelhecimento, ambientais e ocupacionais.

4. Arquitetura para Autenticação Biométrica em Redes Aeronáuticas

A rede utiliza o conceito de conexões segmentadas (i.e. *split connections* [9]). O usuário envia um pacote pela rede interna do avião até o roteador móvel (localizado na aeronave), responsável pelo acesso ao canal de comunicação aeronáutico. Este roteador deve enviar o pacote para a estação aérea-terrestre cujo enlace esteja ativo. O agente intermediário irá encerrar esta conexão TCP e iniciar uma nova conexão utilizando outra socket até o *host* destino. A arquitetura de agentes é descrita na Figura 3 e a interação dos protocolos é ilustrada na Figura 2.

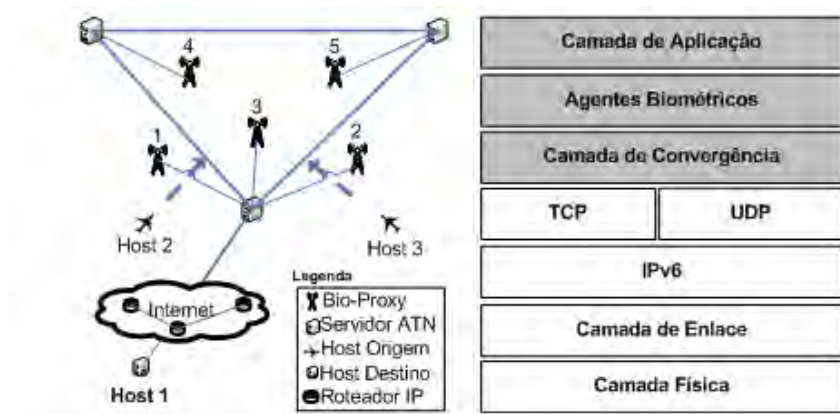


Figura 1. Descrição do cenário aeronáutico e protocolos da arquitetura proposta

4.1. Pilha de Protocolos

A pilha de protocolos é caracterizada pela presença de uma camada *overlay* entre a camada de aplicação e a camada de transporte da pilha TCP/IP, disponibilizando agentes de *software* que realizarão funções de gerenciamento de identidades biométricas. Além dessa camada, deve haver uma camada intermediária (conhecida como adaptador da camada de convergência) responsável pela comunicação entre os protocolos de transporte e a camada *overlay*. A Figura 1(b) ilustra os protocolos e as camadas envolvidas. Os agentes são divididos em quatro tipos: (i) Agente Cliente; (ii) Agente de Captação Biométrica; (iii) Agente Gerenciador; e (iv) Agente Bio-proxy.

O Agente cliente é instalado nas máquinas dos usuários do sistema. Tal agente é responsável pela comunicação com os agentes bio-proxies. Este agente envia informações sobre o seu estado atual na rede, como a associação a uma nova rede. Tal informação será armazenada pelo agente bio-proxy desta nova rede, permitindo que o agente bio-proxy controle os dados biométricos trafegados durante o período de conectividade entre a aeronave e a estação aéreo-terrestre. Além disso, tal agente verifica se a aplicação atual do cliente necessita de dados biométricos, e envia requisições de permissão do uso de um identificador biométrico por um agente bio-proxy, caso seja necessário.

O Agente de Captação Biométrica é instalado em máquinas específicas nos aeroportos. Tal agente realiza a leitura biométrica através de tecnologias biométricas (leitura de impressões digitais, retina, face, veias da mão, etc) e envia tais credenciais biométricas para o agente gerenciador.

O Agente Gerenciador é responsável pelo uso do plano de voo para enviar as credenciais para as respectivas estações aéreo-terrestres. Além disso, o mesmo armazena temporariamente tais credenciais, enviados pelo agente de captação biométrica. A medida que a aeronave vai mudando de estação, as credenciais armazenadas em estações anteriores vão sendo apagadas da memória (*buffer*) destas estações. Devido a velocidade de uma aeronave permanecer na maior parte do tempo constante (aproximando a um modelo matemático de movimento uniforme), o agente pode estimar o tempo de duração que a aeronave permanecerá em cada estação existente no plano de voo. Portanto, o agente pode recalcular caso o plano de voo sofra alterações devido a tempestades ou outros fenômenos não previstos no voo.

O Agente bio-proxy é o agente que armazena as credenciais biométricas quando um plano de voo está sendo realizado. É o agente que se situa nas estações aéreo-terrestres, fazendo o intermédio entre o canal aeronáutico e a rede aeronáutica terrestre. É responsável pela comunicação entre os agentes de mesma natureza. O mesmo se comunica com a aplicação, enviando IDs para o destino quando exigidos.

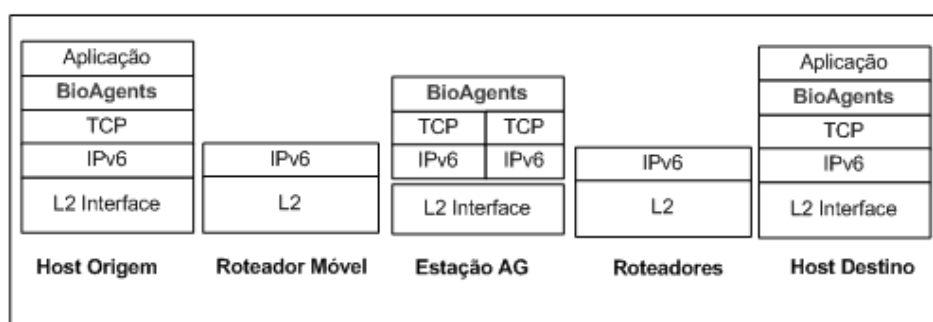


Figura 2. Interação dos Protocolos na Arquitetura Proposta.

O cenário adotado neste trabalho é caracterizado como segue: Em primeiro lugar, cada aeronave possui uma rota definida em um plano de voo elaborado antes da sua decolagem de um determinado aeroporto. O piloto deve segui-lo corretamente, para que não haja imprevistos (acidentes). As estações aéreo-terrestres são definidas por áreas, sendo cada área controlada por um servidor ATN. Estes servidores estão conectados en-

tre si através de enlaces ponto a ponto. Observe que, para redes aeronáuticas maiores, a topologia estrela é sugerida. Em cada rota aeronáutica, a aeronave passará por X estações conhecidas. Um dos servidores atua como gateway, disponibilizando acesso a Internet para toda a rede.



Figura 3. Arquitetura de Agentes de Software do Protocolo Overlay.

5. Estudo de Caso

Antes de viajar, o usuário deve se autenticar em um aeroporto. Sua credencial biométrica será enviada para as estações aéreo-terrestres que se comunicarão com a aeronave durante a rota definida em um plano de voo. Tal informação será armazenada e controlada por agentes intermediários conhecidos como bio-proxies. O usuário não precisará mais se autenticar nas aplicações que exigem autenticação, pois o envio da credencial assim como o seu aceite ou rejeite será negociado pelos agentes intermediários. Caso a identificação seja rejeitada, o agente intermediário repetirá a autenticação até um número de tentativas definido em um parâmetro. Caso ainda haja a falha, o agente intermediário comunicará e passará a função de autenticação para a camada de aplicação do usuário.

5.1. Cenários de Aplicação

A autenticação através de nós intermediários auxilia na viabilidade de diversos tipos de aplicação. Dentre as mais significativas, destacam-se:

- *Acesso a instituições Financeiras:* Sites com serviços financeiros tais como bancos, operadoras de cartão de crédito e cooperativas podem se beneficiar da arquitetura proposta neste trabalho. O usuário não precisaria se autenticar nestes sites a cada nova estação aéreo-terrestre, o que implicaria em redução de fluxo de dados na rede e um atraso menor na comunicação entre hosts.
- *Acesso a sites privados:* Devido ao elevado crescimento do comércio eletrônico nos últimos anos, o usuário poderia efetuar suas compras sem se preocupar com o processo de autenticação, visto que o mesmo estaria temporariamente acessível em cada estação aéreo-terrestre.
- *Acesso a documentos compartilhados:* Sites que compartilham documentos podem utilizar as identidades biométricas sem precisar exigir uma nova captura de identidade cada vez que a aeronave mudar de estação aéreo-terrestre.
- *Jogos para Passageiros:* Aplicações como jogos eletrônicos podem utilizar as identidades biométricas capturadas pelos agentes bio-proxies para permitir a sua

continuidade. Ou seja, os usuários podem jogar entre passageiros de outras aeronaves sem precisar autenticar sua identidade nas estações.

- *Acesso a Intranets*: Empresas podem disponibilizar acesso as suas intranets através do uso das identidades biométricas capturadas nos agentes bio-proxies.

6. Considerações Finais

Este artigo propôs uma arquitetura de rede aeronáutica para oferecer serviços relacionados a segurança biométrica. Para tal, desenvolveu-se uma arquitetura de agentes de *software* que realiza o intermédio entre a aplicação do cliente e o servidor de autenticação biométrica. As estações aéreo-terrestres são responsáveis por este intermédio, armazenando a identidade biométrica durante o voo de uma aeronave convencional que siga um determinado plano de voo. O agente captura as informações do plano de voo, e distribui entre os agentes da rota estimada. Como vantagens de obtenção desta arquitetura, o usuário não se preocupa durante o voo em realizar autenticações biométricas, sendo o agente intermediário responsável por tal atividade. Como atividade futura, pretende-se realizar simulações com simuladores de eventos discretos que investiguem o desempenho desta arquitetura em cenários aeronáuticos.

Referências

- [1] Feighery, P. and Hanson, T. and Lehman, T. and Mondrus, A. and Scott, D. and Signore, T. and Smith, R. and Uhl, G., *The Aeronautical Telecommunications Network (ATN) testbed*, Digital Avionics Systems Conference, 15th Edição, 1996.
- [2] Jahn, A. and Holzbock, M. and Muller, J. and Kebel, R. and de Sanctis, M. and Rogoyski, A. and Trachtman, E. and Franzrahe, O. and Werner, M. and Hu, F., *Evolution of aeronautical communications for personal and multimedia services*, IEEE Communications Magazine, vol. 41, no. 7, pp. 36-43, 2003.
- [3] Schnell, M. and Scalise, S., *NEWSKY Concept for NEtWorking the SKY for Civil Aeronautical Communications*, vol. 22, no 5. IEEE Aerospace and Electronic Systems Magazine, 2007.
- [4] Stephens, B., *Security architecture for aeronautical networks*, The 23rd Digital Avionics Systems Conference, 2004.
- [5] Ehammer, M. and Graupl, T. and Rokitansky, C.H. and Brikey, T., *Security consideration for IP based aeronautical networks*, 27th Digital Avionics Systems Conference, 2008.
- [6] Pankanti, S.; Bolle, R.M.; Jain, A., *Biometrics: The Future of Identification*, IEEE Computer, 2000.
- [7] Jain, A., Hong, L., Pankanti, S., *Biometric Identification*. ACM Communications, vol. 43, no. 2, 2000.
- [8] Fonseca, J., Zanini, A., *Autenticação Biométrica pelas veias da palma da mão*, Instituto Tecnológico de Aeronáutica, 2009.
- [9] Pucha, H. and Hu, Y.C., *Overlay TCP: Ending end-to-end transport for higher throughput*, Poster in ACM SIGCOMM, 2005.
- [10] Brooks, D. and Wilkins, R. and Hoder, D. and InfraStruct, I.G. and LLC, W.C. *Mobile IP communications for aeronautical applications*, The 23rd Digital Avionics Systems Conference, DASC, 2004.

Reconfiguração de Redes de Distribuição de Energia Elétrica Utilizando Metaheurística Colônia de Formigas Modificada

Huilman S. Sanca¹, Niraldo R. Ferreira¹, Benemar A. Souza²

¹Departamento de Engenharia Elétrica – Universidade Federal da Bahia (UFBA)
Rua Aristides Novis, 02, Federação, 40210-630 – Salvador – BA – Brasil

²Departamento de Engenharia Elétrica Universidade Federal de Campina Grande (UFCG)
Av. Aprígio Veloso, 882, Bodocongó, 58429-900 – Campina Grande – PB – Brasil

huilman.sanca@gmail.com, niraldo@ufba.br, benemar@dee.ufcg.edu.br

Abstract. *This paper introduces a modified ant colony optimization algorithm (ACO) to solve the optimal reconfiguration problem of electric power distribution network. This optimization problem is nonlinear, combinatorial. In cooperation, ants find the shortest paths between the nest and a food source by using an indirect mechanism of communication. Seeking to improve the performance of the algorithm we will use a elitism technique in the process of convergence. The method was applied to a 69 buses distribution feeder.*

Keywords. *Reconfiguration, ant colony, elitism, distribution network, active losses, radial load flow.*

Resumo. *Neste artigo apresenta-se um algoritmo colônia de formigas (ACO) modificada para resolver o problema de reconfiguração ótima de redes de distribuição de energia elétrica. Este problema de otimização é não linear e de natureza combinatoria. Em cooperação as formigas descobrem os menores caminhos entre o formigueiro e uma fonte de alimento usando um mecanismo de comunicação indireta. Visando melhorar o desempenho do algoritmo proposto foi dotado de uma técnica de elitismo no processo da convergência. O método foi aplicado a um alimentador de 69 barras.*

Palavras Chaves. *Reconfiguração, colônia de formigas, elitismo, rede de distribuição, perdas ativas, fluxo de carga radial.*

1. Introdução

Os sistemas de distribuição de energia elétrica devem operar de forma confiável e econômica, respeitando os parâmetros de qualidade da energia. O problema da reconfiguração consiste em encontrar outra configuração radial que apresente o menor valor de perdas de potência ativa do sistema melhorando o perfil de tensão nas barras. A reconfiguração de sistemas de distribuição de energia elétrica é levada a efeito operando-se chaves existentes de modo que a topologia da rede seja alterada para que haja transferência de cargas de um alimentador fortemente carregado para outro, relativamente, menos carregado.

Reconfiguração de redes é um problema de otimização combinatoria não diferenciável [Ching-Tzong et al. 2005]. Devido ao elevado número de combinações necessárias, metaheurísticas têm sido empregadas para resolvê-lo. Diferentes métodos

e técnicas de inteligência artificial têm sido aplicados a reconfiguração de redes. Em [Lin et al. 2000], utiliza algoritmos genéticos. Em [Jeon et al. 2002], [Parada et al. 2004], os autores propuseram a aplicação do algoritmo de têmpera simulada (*simulated annealing algorithm*). Em [Olamaei et al. 2007], [Lu et al. 2009], obtém-se a solução do problema mediante o algoritmo de nuvem de partículas (*Particle Swarm Optimization*).

O algoritmo colônia de formigas foi aplicada pela primeira vez ao problema clássico do caixeiro viajante, *Traveling Salesman Problem (TSP)* [Dorigo and Stutzle 2004]. O algoritmo colônia de formigas é inspirado no comportamento das formigas reais, em particular, como é conhecido, as formigas reais são capazes de encontrar o caminho mais curto a partir do formigueiro para a fonte de alimento sem a utilização de sinais visuais, o meio de comunicação das formigas ocorre mediante uma substância química depositada por elas chamada de feromônio.

Neste artigo, se propõe um método ACO modificado que emprega regras de atualização global e local de feromônio, para efeito de melhoramento de seu desempenho. Com a finalidade de demonstrar a eficácia do algoritmo é aplicada para reconfigurar um alimentador de 69 barras apresentado em [Chiang and Jean-Jumeau 1990], e que consta de 69 barras e 73 ligações, fornecendo bons resultados.

2. Formulação do Problema

2.1. Formulação do problema de reconfiguração de redes para o ACO Elitista

A reconfiguração de redes pode ser formulado como um problema de otimização não linear mista, de variáveis inteiras e reais, cuja expressão é a seguinte:

$$\text{Minimize } f(x) = P_{T,Perdas} = \sum_{i=1}^{N_B-1} P_{perdas}(i, i+1) \quad (1)$$

Sujeito as restrições:

1. de fluxo de carga;
2. limites de corrente nos trechos:

$$|I_j| \leq I_{j,max}; \forall j, j \in N_T \quad (2)$$

3. configuração radial da rede.

$P_{T,perdas}$ é a perda de potência total do sistema, $|I_i|$ e $I_{i,max}$ são a amplitude da corrente e o limite máximo de corrente em cada trecho i , respectivamente, N_B é o número de barras e N_T é o número de trechos do sistema.

O problema de otimização com restrições expresso em (1), pode ser convertido num problema de otimização irrestrito cuja função objetivo incorpore a restrição de corrente máxima nos trechos [Lin et al. 2000].

$$\text{Minimize } f(x) = \sum_{i=1}^{N_B-1} P_{perdas}(i, i+1) + \sum_{j=1}^{N_T} \lambda_I (I_j - I_{j,max})^2 \quad (3)$$

sendo λ_I um fator de penalidade com respeito à corrente admissível no trecho.

2.2. Cálculo das perdas

O cálculo das perdas de potência é parte do problema do fluxo de carga, que aqui é resolvido empregando-se o método da soma de potência (MSP). O MSP, é um método iterativo nas variáveis de perdas de potência ativa e reativa do tipo *forward - backward* [Haque 2000]. Os valores absolutos das tensões das barras são calculadas sequencialmente no sentido das subestações para as barras terminais (*forward*). As potências dos trechos são calculados sequencialmente no sentido das barras terminais para as subestações (*backward*). Para o cálculo de um trecho ($i, i+1$). Referido à Figura 1, podem-se escrever estas equações como em [Haque 2000]:

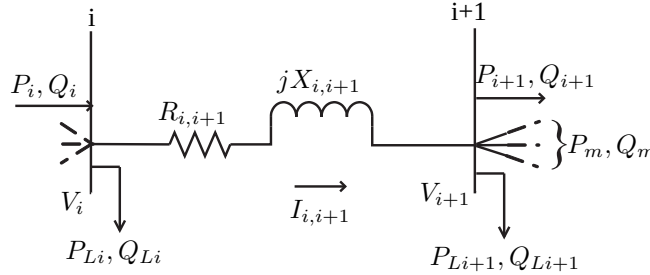


Figura 1. Trecho de uma rede de distribuição simples.

$$P_i = P_{i+1} + P_{Li+1} + R_{i,i+1} \left[\frac{P_i^2 + Q_i^2}{|V_i|^2} \right] \quad (4)$$

$$Q_i = Q_{i+1} + Q_{Li+1} + X_{i,i+1} \left[\frac{P_i^2 + Q_i^2}{|V_i|^2} \right] \quad (5)$$

$$|V_{i+1}|^2 = |V_i|^2 - 2(R_{i,i+1}P_i + X_{i,i+1}Q_i) + (R_{i,i+1}^2 + X_{i,i+1}^2) \frac{(P_i^2 + Q_i^2)}{|V_i|^2} \quad (6)$$

sendo P_i e Q_i são os fluxos de potência ativa e reativa no trecho entre as barras i e $i+1$, respectivamente. P_{Li} e Q_{Li} são as potências ativa e reativa das cargas instaladas na barra i , respectivamente. A resistência e a reatância do trecho entre as barras i e $i+1$, são indicadas por $R_{i,i+1}$ e $X_{i,i+1}$, respectivamente.

Uma expressão alternativa à (6) é:

$$|V_{i+1}| = \sqrt{A + \sqrt{A^2 - B}}, \quad (7)$$

sendo:

$$A = \frac{|V_i|^2}{2} - (R_{i,i+1}P_i + X_{i,i+1}Q_i)$$

e

$$B = (R_{i,i+1}^2 + X_{i,i+1}^2)(P_i^2 + Q_i^2).$$

A perda de potência ativa no trecho é dada por:

$$P_{perdas}(i, i+1) = R_{i,i+1} \frac{P_i^2 + Q_i^2}{|V_i|^2}. \quad (8)$$

A perda de potência ativa total do sistema $P_{T,Perdas}$, é a soma das perdas de potência ativa de todos os trechos do sistema:

$$P_{T,Perdas} = \sum_{i=1}^{N_B-1} P_{perdas}(i, i+1). \quad (9)$$

3. Algoritmo Colônia de Formigas

3.1. O comportamento da colônia de formigas

O algoritmo colônia de formigas, ACO, é uma metaheurística inspirada nas formigas reais [Dorigo and Stutzle 2004], [Ding and Loparo 2012].

Na busca por alimento as formigas inicialmente se movimentam sem orientação tomando decisões com base em critérios individuais. Mais tarde, as formigas que escolheram o caminho mais curto entre o formigueiro e a fonte de alimento completarão suas expedições mais rapidamente. Isto fará que mais formigas escolham o menor caminho devido à maior concentração de feromônio. No fim, todas as formigas irão escolher o mesmo caminho devido à atualização constante do feromônio. Este comportamento mostra o paradigma fundamental do algoritmo de busca por colônia de formigas.

3.2. Escolha pseudo-aleatória das ligações

Trazendo o modo natural das formigas descobrir e coletar alimentos para resolver o problema da reconfiguração de redes se pode imaginar o seguinte: a formiga escolhe a barra seguinte a ser visitada com base em seu próprio conhecimento que tem da rede (resistência das ligações entre a barra onde ficou a formiga e as barras vizinhas a ela) e no conhecimento coletivo (quantidade de feromônio depositado nas ligações durante o percurso). O conhecimento coletivo é cumulativo, sendo alterado sempre que uma nova configuração radial se completa. A probabilidade de uma das ligações vizinhas vir a ser escolhida por uma formiga é dada pela expressão:

$$Prob_{(i,j)} = \begin{cases} \frac{[\tau_{(i,j)}]^\alpha [\eta_{(i,j)}]^\beta}{\sum_{m \in J_{K(i)}} [\tau_{(i,m)}]^\alpha [\eta_{(i,m)}]^\beta} & \text{se } \forall j \in J_{K(i)}; \\ 0 & \text{se } \forall j \notin J_{K(i)}, \end{cases} \quad (10)$$

sendo τ a quantidade de feromônio na ligação escolhida (i, j) , cuja resistência é $1/\eta_{(i,m)}$. α e β são os pesos do grau de atratividade e a visibilidade do feromônio respectivamente. $J_{K(i)}$ é o conjunto das ligações vizinhas que poderão ser visitadas pela formiga K que se encontra no nó i .

3.3. Regra de atualização local do feromônio

Logo após as formigas percorrerem cada trecho, entre as barras i e j , aplica-se a regra de atualização local do feromônio, conforme a seguinte expressão:

$$\tau_{(i,j)} = (1 - \rho)\tau_{(i,j)} + \rho\tau_0 \quad (11)$$

Onde $\tau_{(i,j)}$ é a carga de feromônio na ligação (i, j) , τ_0 é o valor inicial do feromônio e ρ é a taxa de evaporação do feromônio, $\rho \in [0, 1]$. Esta regra de atualização local serve para diminuir a possibilidade de estagnação do processo e aumentar a probabilidade de encontrar uma melhor solução até o momento [Dorigo and Stutzle 2004].

3.4. Regra de atualização global do feromônio

Logo após uma configuração radial factível ser achada são calculadas as perdas de potência ativa da rede e a carga de feromônio nas ligações da melhor configuração é incrementada do seguinte modo:

$$\tau_{(i,j)} = (1 - \sigma)\tau_{(i,j)} + \sigma\Delta\tau \quad (12)$$

sendo:

$$\Delta\tau = \frac{1}{F_{melhor}} \quad (13)$$

e F_{melhor} é o valor da função objetivo da melhor solução achada até o momento, σ é a taxa de evaporação da regra de atualização global, $\sigma \in [0, 1]$. O feromônio é uma substância volátil o suficiente para evitar que soluções antigas sejam demasiadamente persistentes. $(1 - \sigma)$ modela a evaporação do feromônio.

4. Aplicação do Algoritmo para o Problema de Reconfiguração de Redes

Inicialmente todos os trechos têm a mesma quantidade de feromônio, todos os nós-fontes estão ligados, portanto os nós de carga estão todos desligados, mas nenhuma ligação está ativada. Nó ligado é aquele onde está a formiga no momento, ligação ativada é aquela que já foi percorrida e ligação ativável é aquela ligação vizinha para a qual a formiga pode se deslocar. O algoritmo *ACO Elitista* proposto é aquele que se apresenta na Figura 2.

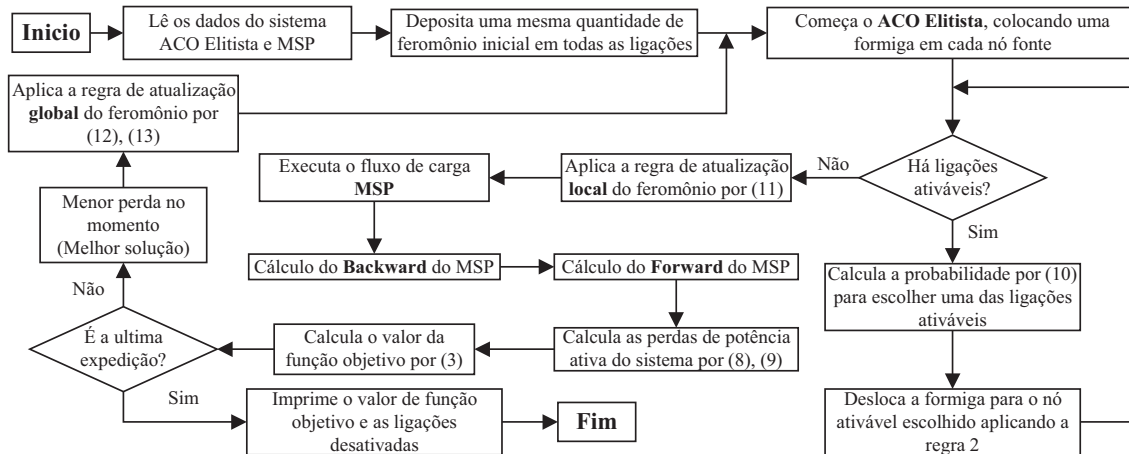


Figura 2. Fluxograma do ACO com Elitismo.

As formigas partem simultaneamente de forma aleatória de nós ligados respeitando as seguintes regras:

1. As formigas se deslocam exclusivamente por ligações ativáveis;
2. Quando uma formiga chega ao nó desligado da ligação ativável que tenha percorrido:
 - este nó torna-se ligado e a ligação ativada;
 - surge outra formiga para ocupar o nó originalmente ligado deixado por ela;

3. O percurso de uma formiga se completa quando ela não puder mais seguir por ligações ativáveis;
4. A expedição termina quando nenhuma formiga tiver mais mobilidade, ou seja, quando não houver nenhuma ligação ativável.

Durante uma iteração, apenas uma lista de barras a serem visitadas (barras ativáveis vizinhas ao nó ligado) é utilizada e escolhida de forma probabilística, atualiza-se a lista conforme os agentes caminham sobre a rede. Quando um agente se encontra sobre a barra i , seleciona a barra vizinha j (diretamente conectada a i) a ser visitada com base na concentração de feromônio sobre a linha (i,j) e no inverso da resistência desta, só são consideradas barras ativáveis aquelas barras que são diretamente conectadas a barra ligada (barra onde a formiga esta no momento).

5. Teste e Resultados

O algoritmo proposto foi implementado em *Matlab*® e testado com o alimentador de 69 barras da Figura 3, no qual há 5 chaves de interconexão e 73 chaves seccionadoras, originalmente 68 chaves estão fechadas (chaves de 1 a 68) e 5 abertas (chaves de 69 a 73). Para esta configuração inicial as perdas ativas iniciais são de 20,9 kW e a tensão de linha na saída da subestação (barra 1) é de 12,66 kV. Os parâmetros utilizados no algoritmo são mostrados na Tabela 1. A solução, que se vê na tabela 2, foi alcançada em 3,79 segundos utilizando-se um computador com processador *Intel core i5-2410M* de 2,3 GHz e 6 GB de RAM.

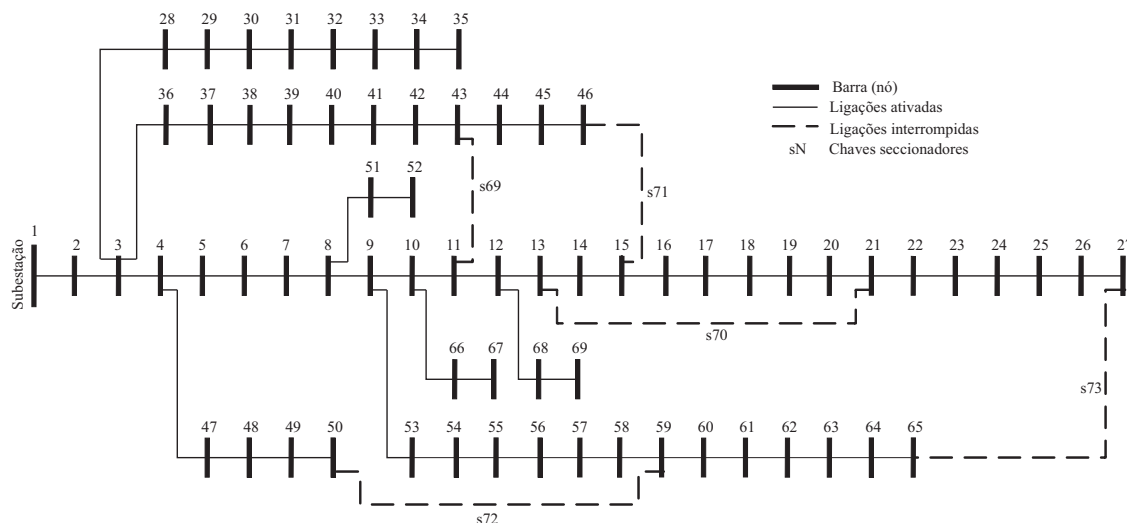


Figura 3. Sistema de distribuição de 69 barras e 73 ligações, [Chiang and Jean-Jumeau 1990].

A perda ativa total para a configuração inicial do sistema é de 20,9 kW, enquanto para a configuração obtida pelo algoritmo proposto elas foram de 9,35 kW, isto equivale a uma redução de 55,26%.

Fazendo-se uma comparação com os resultados publicados por [Chiang and Jean-Jumeau 1990] e [Abdelaziz et al. 2012], tal como mostrado na

Tabela 1. Parâmetros usados para o teste do algoritmo

Parâmetros	Símbolo	Valor	Parâmetros	Símbolo	Valor
Alfa	α	0,1	Expedições		170
Beta	β	1,9	Sigma	σ	0,01
Rô	ρ	0,1	Feromônio inicial	τ_0	1
Tolerância do MSP	ε	10^{-3}			

Tabela 2. Resultados do teste

Configuração	Perdas finais (kW)	Redução (%)	Chaves desligadas
Inicial	20,9	- - -	s69, s70, s71, s72, s73
ACO Elitismo	9,35	55,26	s10, s57, s62, s70, s71
[Chiang and Jean-Jumeau 1990]	9,34	55,31	s15, s57, s62, s70, s71
[Abdelaziz et al. 2012]	9,355	54,30	s14, s58, s61, s69, s70

tabela 2, nota-se o bom desempenho do algoritmo proposto. A Figura 4 apresenta uma comparação das tensões em cada barra nas configurações inicial e final (ótima).

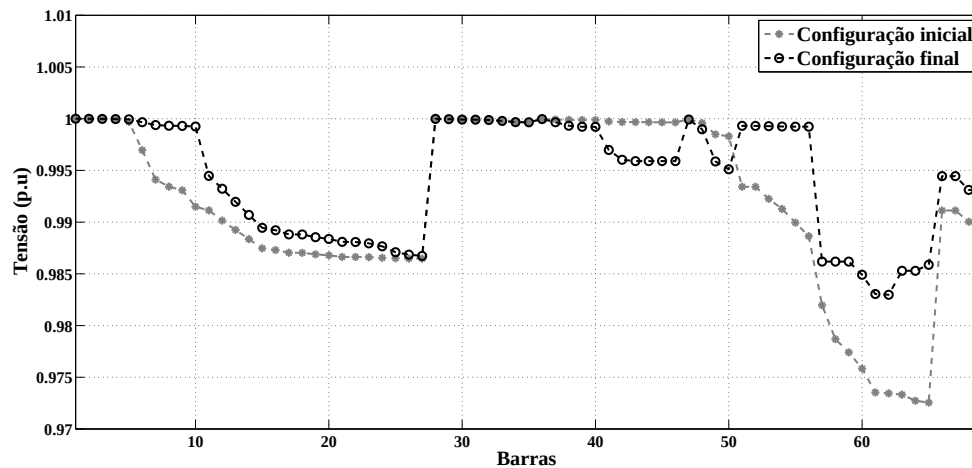


Figura 4. Comparação de tensões das barras para a configuração inicial e a configuração ótima encontrada.

Observa-se a melhoria obtida no perfil de tensão decorrente da reconfiguração alcançada desativando-se as chaves seccionadoras s10, s57, s62, s70, s71 que equivalem aos trechos dos nós: 10-11, 57-58, 62-63, 13-21, 15-46.

6. Conclusão

Um algoritmo de formigas modificado (ACO Elitista) foi proposto neste artigo com o objetivo de encontrar uma configuração radial ótima, minimizando as perdas de potência ativa totais. O escopo do trabalho foi resolver o problema de reconfiguração ótima de redes de distribuição de energia elétrica aplicando uma metodologia inspirada no comportamento de colônia de formigas.

Para melhorar o desempenho, o algoritmo desenvolvido incorpora uma técnica de elitismo no processo de convergência mediante regras de atualização local e global do feromônio. Dos resultados obtidos, conclui-se que o ACO proposto é promissor para a solução do problema de reconfiguração de redes apresentando um bom desempenho. Pretende-se agora investigar melhorias nas propriedades de convergência do algoritmo.

7. Agradecimentos

Os autores deste artigo agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES e a Fundação de Amparo à Pesquisa do Estado da Bahia - FAPESB.

Referências

- Abdelaziz, A. Y., Osama, R. A., and El-khodary, S. M. (2012). Reconfiguration of distribution system for loss reduction using the hyper-cube ant colony optimization algorithm. *IET Generation, Transmission and Distribution*, 6(2):176–187.
- Chiang, H.-D. and Jean-Jumeau, R. (1990). Optimal network reconfigurations in distribution systems: part 2 : Solution algorithms and numerical results. *IEEE Transactions on Power Delivery*, 5(3):1568–1574.
- Ching-Tzong, S., Chung-Fu, C., and Ji-Pyng, C. (2005). Distribution network reconfiguration for loss reduction by ant colony search algorithm. *Electric Power Systems Research*, 7:190–199.
- Ding, F. and Loparo, K. A. (2012). A simple heuristic method for smart distribution system reconfiguration. *Energytech, IEEE Conference*, pages 1–6.
- Dorigo, M. and Stutzle, T. (2004). Ant colony optimization. *MIT Press*, 3(2):12.
- Haque, M. H. (2000). A general load flow method for distribution systems. *Electric Power System Research*, 54(1):47–54.
- Jeon, Y., Kim, J. C., Kim, J., Shin, J. R., and Lee, K. (2002). An efficient simulated annealing algorithm for network reconfiguration in large-scale distribution systems. *IEEE Transactions on Power Delivery*, 17:1070–1078.
- Lin, W. M., Cheng, F. S., and Tsay, M. T. (2000). Distribution feeder reconfiguration with refined genetic algorithm. *IEE Proceedings Generation, Transmission and Distribution*, 147(6):349–354.
- Lu, L., Liu, J., and Wang, J. (2009). A distributed hierarchical structure optimization algorithm based poly-partical swarm for reconfiguration of distribution network. *SUPERGEN 09 International Conference on Sustainable Power Generation and Supply*, pages 1–5.
- Olamaei, J., G., G., and T., N. (2007). An approach based on particle swarm optimization for distribution feeder reconfiguration considering distributed generators. *Power Systems Conference: Advanced Metering, Protection, Control, Communication, and Distributed Resources*, pages 326–330.
- Parada, V., Ferland, J. A., Arias, M., and Daniels, K. (2004). Optimization of electrical distribution feeders using simulated annealing. *IEEE Transactions on Power Delivery*, 19:1135–1141.

Uma proposta de desenvolvimento de um sistema de software para apoio ao gerenciamento de doenças relacionadas à Leishmaniose

Ludmylla Lopes Lobato¹, Luiz Arthur Lopes Lobato¹, Luanna Lopes Lobato²

¹Curso de Medicina – Universidade Presidente Antônio Carlos (UNIPAC)
Juiz de Fora – MG – Brasil

²Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Catalão – GO – Brasil

{ludmyllalopeslobato, luizarthurlopeslobato}@gmail.com,
l11@catalao.ufg.br

Abstract. *The American Cutaneous Leishmaniasis, or Leishmaniasis, is an endemic zoonosis in the tropics and neotropics caused by the protozoan parasites from the Leishmania, which was identified with more 20 different species related to it. The disease is found worldwide and, according to published studies, present a high incidence in the Brazil, with six species associated that can be contracted by humans. Thus, this work is focused on present a study about Leishmaniasis, in order to propose the development of a software system to facilitate the identification of this disease, as well as facilitate the verification of its impacts and provide the monitoring of the disease and possible ways to prevention. In this sense, we presented an initial study related to the disease in a region of Brazil, followed by the proposal of a software system to help manage the data collected through the observance of the disease type.*

Resumo. *A Leishmaniose Tegumentar Americana, ou simplesmente Leishmaniose, é uma zoonose endêmica nos trópicos e neotrópicos causada pelos protozoários parasitas do gênero Leishmania, possuindo mais de 20 diferentes espécies diferentes. A doença é encontrada em todo o mundo e, de acordo com estudos apresentados, no Brasil apresenta alta incidência, sendo seis espécies associadas à doença nos seres humanos. Assim, este trabalho tem como foco o estudo da Leishmaniose e a proposta de um sistema de software para identificação da doença, verificação de seus impactos, bem como o monitoramento da doença e a possível prevenção. Neste sentido, apresenta-se um estudo inicial sobre a doença em uma região do Brasil, seguido pela proposta de um sistema de software para auxiliar no gerenciamento dos dados coletados relacionados à observância da doença.*

1. Introdução

A ciência tem avançado muito em relação a identificação de doenças e seus possíveis tratamentos, sendo interligadas várias área de conhecimento para que as melhores soluções sejam produzidas, como a química, possibilitando a intervenção por diferentes

medicamentos, a biologia, em que são apresentadas as próprias defesas do corpo humano, a computação permitindo o uso de softwares para auxiliar na identificação e acompanhamento das doenças. De uma maneira particular, o uso de sistemas desenvolvidos com o propósito de, além de identificar, monitorar as doenças e seus avanços tem sido frequentemente apresentado em vários estudos.

Neste contexto, este trabalho é justificado visto que o desenvolvimento de software é um dos grandes aliados da medicina, uma vez que possibilita e facilita a execução automatizada de estudos que poderiam demorar dias ou até anos. Assim, apresenta-se, neste artigo, um estudo inicial sobre os reflexos da leishmaniose no Brasil e a possível redução do impacto dessa doença e controle da mesma, se constatada e, possivelmente, acompanhada através da otimização de sistemas de software.

A leishmaniose tegumentar americana (LTA) é uma zoonose endêmica nos trópicos e neotrópicos causada pelos protozoários parasitas do gênero *Leishmania* (Krauss *et al.*, 2003) que possuem mais de 20 espécies diferentes. A doença é encontrada em todo o mundo, com uma prevalência estimada de 12 milhões de casos por ano, havendo cerca de 50.000 mortes, e uma estimativa de que 350 milhões de pessoas em todo o mundo estão em situação de risco de adquirir a doença.

No Brasil, pode ser encontrada em todas as macrorregiões, com alta incidência anual, estando seis espécies associadas à doença no homem, sendo uma do gênero *Leishmania* e cinco do subgênero *Viannia*. A *Leishmania* (*Viannia*) *brasiliensis* caracteriza-se por diversas manifestações clínicas, que abrangem desde pequenos nódulos à destruição da mucosa, sendo a principal espécie responsável pela LTA no país, sendo o foco de estudo desta pesquisa inicial.

O trabalho é apresentado como segue: Na Seção 2 é apresentada uma pesquisa relacionada. Na Seção 3 são descritas informações sobre a Leishmaniose, sendo esta a doença em estudo neste artigo. A proposta do sistema de software é apresentada na Seção 4. Já na Seção 5 são apresentadas as conclusões, sendo apresentadas algumas discussões sobre o tema abordado, bem como os impedimentos para esta pesquisa e sugestões de trabalhos futuros.

2. Trabalho Relacionado

Como exemplo de trabalho relacionado, tem-se o software VINA, proposto pelo “The Scripps Research Institute”, situado em La Jolla, California. Este software é usado para executar experimentos virtuais com o objetivo de encontrar milhões de componentes que podem ser capazes de desativar algumas proteínas específicas, as quais são essenciais para a sobrevivência de parasitas (World Community Grid).

Neste sentido, a pesquisa apresenta que uma busca, ou também conhecida como triagem, para obter os melhores compostos de drogas potenciais é um passo inicial no processo de desenvolvimento de tratamentos eficazes para a doença. Ainda é ressaltado que, com suficiente poder computacional esta rastreio pode ser feita muito mais rapidamente do que através de experimentos de laboratório convencionais.

No entanto, se os meios computacionais não são utilizados em sua alta performance, ainda seriam necessários cerca de 120 anos para os pesquisadores realizarem tal triagem. Um exemplo de sistema que pode auxiliar em muito tal tarefa é o

uso de Grids computacionais¹. O poder da “World Community Grid” pode reduzir o tempo necessário para a menos de um ano, facilitando dessa forma o trabalho dos pesquisadores e permitindo que os melhores candidatos identificados por este projeto possam levar ao desenvolvimento de melhores medicamentos para combater a leishmaniose.

3. A Leishmaniose

A leishmaniose tegumentar americana (LTA) é uma doença infecciosa, não contagiosa, que provoca a destruição das células fagocitárias pelos parasitas amastigotas da leishmania. Tal doença pode levar a um estado clínico conduzido a um processo de lesões dérmicas e da mucosa, o que caracteriza-se pelas formas localizada e disseminada, sendo um dos primeiros sinais clínicos apresentados pelo hospedeiro. Ela pode manifestar-se com maior ou menor severidade e até mesmo passar despercebida, pois tudo depende da patogenicidade do agente causador e suscetibilidade do hospedeiro.

A contaminação humana ocorre através da picada do vetor mecânico flebotômico fêmea infectado, no ². Como meio de inocular seu sangue, o inseto lacera a pele do ser humano, transmitindo assim para o hospedeiro a forma promastigota do parasita. Em contrapartida, a contaminação do vetor acontece por meio de vertebrados infectados, o que pode ocorrer durante a inoculação do sangue contaminado com a forma amastigota pelo inseto (Krauss *et al.*, 2003).

Em relação ao habitat dos flebótomos, estes costumam viver em áreas quentes e úmidas,. Como exemplo pode-se citar troncos de árvores podres e folhas, os quais são ideais para o desenvolvimento do mosquito, que, uma vez estando em forma de larvas, se alimentam de materiais orgânicos.

O mosquito flebótomo pica entre o entardecer e a noite. Segundo estudos feitos, os dados epidemiológicos da microrregião de Juiz de Fora – Minas Gerais, local onde a pesquisa têm sido conduzida, a maior incidência ocorre por volta das oito e dez horas da noite. No entanto, se o ser humano tem contato direto com o habitat natural, o mosquito pode picar a qualquer momento.

3.1. Morfologia e Ciclo Biológico

Os protozoários de gênero *Leishmania* são parasitas digenéticos, causadores das doenças conhecidas por Leishmanioses (Krauss *et al.*, 2003). Este é um protozoário dimórfico, que existe como promastigota flagelado no tubo digestivo de sugadores de sangue invertebrados (fêmeas de flebotômicos), de corpo alongado, medindo entre 10 e 15 micras de longitude, com núcleo na parte média do corpo e no externo anterior do organismo está o cinetoplasto onde sai o flagelo para sua locomoção, sendo a forma que inocula os vertebrados. E como amastigota, intracelular obrigatório, medindo de 2 a 5 micras de longitude, sem flagelo, arredondada, localiza-se no interior de macrófagos e

¹ Grids são uma forma de computação distribuída, em que um "super computador virtual" é composto de muitos computadores em rede de baixo acoplamento, agindo em conjunto para realizar tarefas de grande porte.

² Explicação de termos em: *Etymological Dictionary of Veterinary Medicine and its Auxilliary Sciences*

células dendríticas de hospedeiros vertebrados. Podendo também infectar fibroblastos e neutrófilos. No citoplasma observa-se o núcleo que pode ser ovóide ou esférico situado na parte central da célula.

A reprodução ocorre de forma assexuada, caracterizando-se por divisão binária longitudinal, ocorrendo de forma metacíclica. O ciclo de vida ocorre em dois tipos de hospedeiros. Nos hospedeiros invertebrados é iniciado quando o flebótomo fêmea pica a pele de vertebrados infectados, formando uma lesão com sangue e macrófagos na derme onde estão os parasitas. Durante esse repasto sanguíneo se ingere as células fagocíticas, principalmente macrófagos, abrigando essas formas amastigotas. Após o rompimento dos macrófagos, os amastigotas no intestino do inseto assumem a forma de promastigotas procíclicos, que se multiplicam ativamente por divisão binária e estão aderidos a parede do intestino do inseto evitando sua expulsão. Tendo inoculado ao corpo do vetor mecânico, os promastigotas passam por alguns processos para infectar o hospedeiro. Detalhes deste ciclo podem ser verificadas em (Krauss *et al.*, 2003).

3.2. Manifestações Clínicas

A picada do flebótomo em algumas vezes pode ser verificada pelo surgimento de lesão na pele. A forma cutânea é a representação clínica mais comum, podendo surgir como pápulas e se desenvolver úlceras. A forma mucosa, na maioria das vezes são secundária às lesões cutâneas, surgindo geralmente meses ou anos após a resolução das lesões de pele.. Em geral acometem as vias aéreas superiores, como cavidades nasais (principalmente o septo nasal), seguidas da faringe, laringe e cavidade oral. As queixas mais comuns no acometimento nasal são: obstrução, epistaxes, rinorréia e crostas; da faringe, odinofagia; da laringe, rouquidão e tosse; da cavidade oral, feridas na boca. As lesões podem ser discretas, sendo apresentados poucos sintomas, por isso a necessidade de se buscar a identificação de doença em mucosas, uma vez que as lesões mucosas não curam espontaneamente. Através do exame clínico, é possível observar infiltração, ulceração, perfuração do septo nasal, lesões ulcero vegetantes, ulcero crostosas ou ulcero destrutivas.

Dessa forma, verificou-se a importância em reportar em um sistema de software os dados coletados sobre a doença e seus impactos no paciente. Tal sistema tem por objetivo permitir a otimização na busca pelos resultados, bem como possibilitar a viabilidade de se associar os medicamentos ao estágio em que a doença se encontra. Assim, espera-se com o sistema permitir a verificação, análise e acompanhamento do prognóstico do paciente de forma ágil e precisa.

4. Proposta de Trabalho

As formas de diagnóstico para esta doença estão presentes em âmbito clínico, epidemiológico e laboratorial. Visto que a LTA é um problema recorrente em várias partes do mundo e se enquadra em um problema endêmico de saúde no Brasil. Dessa forma, justifica-se a necessidade de um diagnóstico eficiente e eficaz, assim espera-se ter um diagnóstico rápido e preciso que identifique os indivíduos infectados e exponha resultados sobre a existência de material do parasita em indivíduos considerados curados, tais ações devem ser verificadas de modo a prevenir reincidentes, prover

acompanhamento dos casos já identificados, bem como monitorar e controlar até a expansão da endemia.

O tratamento da doença é através de fármacos e o controle da infecção Leishmaniose em pacientes pode ser verificada devido à resposta imune celular, de modo que a reativação da doença pode ocorrer como resultado de imunossupressão.

Como meio de facilitar a identificação e acompanhar o tratamento da doença, propõe-se neste trabalho o desenvolvimento de um sistema de software. Tal sistema deve ser integrado com diferentes projetos, em que tem-se como entrada a possibilidade de identificação de casos de pacientes infectados, os quais são mantidos através da base de dados possibilitando o acompanhamento e, eventual, tratamento desses casos.

Neste sentido, para a proposta deste trabalho inicialmente desenvolveu-se uma pesquisa sobre o estudo da arte da Leishmaniose no Brasil. Como exemplo prático, estudos de casos foram verificados na cidade de Juiz de Fora, Minas Gerais, onde os autores deste artigo puderam realizar pesquisas relacionadas ao contexto da Leishmaniose na cidade.

Esta pesquisa está sendo desenvolvida por pesquisadores de duas instituições de ensino, em que um estudo inicial já foi iniciado com vistas a atender os objetivos descritos neste trabalho. Assim, a escolha pelo estudo inicial sobre o cenário da Leishmaniose na cidade de Juiz de Fora, Minas Gerais, é justificada devido a disponibilidade de recursos providos pela Universidade Presidente Antônio Carlos (UNIPAC), através do auxílio de professores e alunos do curso de Medicina. Adicionalmente, este trabalho tem como apoio a Universidade Federal de Goiás, campus Catalão, mais especificamente o grupo de pesquisa ReSEArch (*Reuse in Software Engineering and Architecture*), o qual contribui para os estudos referentes ao desenvolvimento do sistema de software para facilitar o gerenciamento dos dados coletados.

Em relação ao desenvolvimento do sistema, cinco reuniões foram realizadas entre os pesquisadores envolvidos de modo a detalhar as necessidades e os requisitos funcionais e não funcionais que o sistema deva prover. Para tanto, utilizar-se-á a Engenharia de Software como metodologia para desenvolvimento do sistema, seguindo as recomendações impostas por Pressman (2001). Dessa forma, os requisitos, casos de uso, diagrama de classe, bem como diagramas de atividades foram desenvolvidos e validados por outros integrantes do grupo de pesquisa ReSEArch e verificados pelos estudantes de medicina envolvidos.

5. Conclusão

A Leishmaniose *brasiliensis* é uma doença hemoprotozoária, sendo transmitida pela picada do inseto flebótomo para o homem e outros animais como, equinos, cães e roedores (Krauss et al., 2003). É uma doença que agride as áreas primárias de entrada ao organismo, como a pele e a mucosa. Os sintomas mais comuns são da doença estão relacionados ao aparecimento de lesões epidérmicas em forma de ulcerações únicas, múltiplas, disseminadas ou difusas, com contornos e inodoros, e, em algumas situações, é também apresentada febre. Pode-se citar algumas áreas específicas em que são

observadas os impactos da doença, como lábio superior, nariz, região malar, pavilhão auricular, pernas, pés, braços e antebraços.

Ao ser detectada a infecção, o sistema imune hospedeiro é ativado e passa a responder positivamente na tentativa de eliminar o parasito. No entanto, a *Leishmania* consegue evadir aos eventos do organismo e explorar seu sistema imune, modulando-o em benefício próprio, para assim formar um ambiente favorável ao estabelecimento da infecção. Esse desequilíbrio na relação parasita-hospedeiro que causa a eliminação da infecção.

Um dos grandes problemas relacionados a esta doença é que o diagnóstico não é simples, devido à demora de manifestação do quadro infeccioso o que pode chegar de meses a anos. Assim, pode ser tornar difícil uma identificação diagnóstica, principalmente casos fora das áreas endêmicas.

Com o desenvolvimento deste trabalho, mesmo que inicial, foi possível identificar informações relevantes sobre a *Leishmaniose brasiliensis* (*Leishmania Tegumentar Americana*), o que fortaleceu o conhecimento da equipe de pesquisa em relação aos impactos da doença e possíveis danos irreparáveis. Espera-se que uma pesquisa mais detalhada possa respaldar-nos o suficiente para contribuir para a proposta de desenvolvimento de um sistema de software que possa auxiliar no tratamento da doença descrita.

Neste sentido, como trabalho futuro propõe-se que mais estudos sejam integrados. Assim, sugere-se que seja feita uma pesquisa sistemática sobre as pesquisas que têm sido publicadas sobre *Leishmaniose* no mundo, de modo a identificar os principais pontos que foram observados por estes trabalhos. Tais resultados são relevantes visto que podem ser utilizados para desenvolvimento e melhoria do sistema proposto neste artigo.

Adicionalmente, espera-se desenvolver mais pesquisas relacionando a medicina às facilidades providas pela otimização da execução das tarefas através da Computação. O alinhamento e parceria entre essas áreas têm se mostrado interessantes e essenciais, uma vez que podem facilitar a identificação, prover o tratamento, facilitar o acompanhamento e a prevenção de doenças inerentes ao ser humano.

Referências

- Pressman, R. S. 2001. *Software Engineering: A Practitioner's Approach* (5th ed.). McGraw-Hill Higher Education.
- Krauss, H., Weber, A., Appel, M., Enders, B., Graevenitz, A. V., Isenberg, H. D., Schiefer, H.G., Slenczka, W. and Zahner, H. 2003. *Zoonosis. Infectious Diseases Transmissible from Animals to Humans*. ASM Press. American Society for Microbiology, Washington DC., USA. 2003. 3rd Edition, 456 pages. ISBN: 1-55581-p. 236-8.
- World Community Grid – Technology solving problems. “Drug Search for Leishmaniasis”, Disponível em: <http://www.worldcommunitygrid.org/research/dsfl/overview.do>. Acesso em: Fevereiro de 2013.

Um breve estudo sobre Gerenciamento de Riscos em Projetos de *Software* no Brasil

Lorena T. de Carvalho¹, Hugo S. de Andrade², Luanna L. Lobato¹

¹Departamento de Ciência da Computação – Universidade Federal de Goiás, Campus Catalão (UFG/CAC) - Sala 202 - Bloco Didático 2, Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão/GO CEP: 75704-020.

²Departmento de Ciência da Computação – Universidade Federal da Bahia (UFBA) – Salvador, BA - Brasil

{lorena.computacao, hugosica, luannalopeslobato}@gmail.com

Abstract. *This paper presents a systematic study on the state of the art of risk management during software development. The studies considered in this work were made available through scientific events held in Brazil, and also obtained from journals and the internet. The goal is to collect information about what has been developed in order to verify how risk management has collaborated over the lifecycle of software projects and proposal for further development of studies on this area.*

Resumo. *Neste artigo é apresentado um estudo sistemático sobre o estado da arte de gerenciamento de riscos durante o desenvolvimento de software. Como fonte de pesquisa são considerados trabalhos que foram disponibilizados em eventos científicos realizados no Brasil, sendo verificados eventos relacionados à área, revistas científicas e a internet. O objetivo é coletar informações sobre o que já foi desenvolvido a fim de verificar o quanto o gerenciamento de riscos tem colaborado em relação ao ciclo de vida dos projetos de software para posterior proposta e desenvolvimento de estudos sobre esta área.*

1. Introdução

O crescente uso de tecnologia para o desenvolvimento e prestação de serviços têm exigido da indústria de *software* projetos cada vez mais complexos. Para tanto, é normal que, em projetos de *software*, surjam problemas relacionados à qualidade, cronograma, custos, dentre outros. Tais situações podem estar relacionadas a riscos que não foram previamente identificados ou, simplesmente, foram ignorados durante o planejamento do projeto.

Desde 1995 tais problemas tem sido discutidos em estudos da engenharia de *software*. Segundo Pressman (1995), a utilização de modelos de melhorias, processos, métodos e ferramentas vem para aumentar a probabilidade de sucesso na execução de projetos de *software*, focando na qualidade de seus produtos. Diante disso, à medida que a complexidade dos *softwares* aumenta, cresce também a necessidade de metodologias para gerir as incertezas contidas nos projetos.

Assim, neste trabalho, tem-se como justificativa a necessidade de verificar o cenário atual em relação ao que tem sido desenvolvido sobre gerenciamento de riscos no Brasil, visto que não foi identificado nenhum estudo sistemático que contemple

discussões ou experiências de projetos brasileiros. Dessa forma, espera-se com este trabalho reunir dados através de estudos sobre gerenciamento de riscos para futuras propostas de melhorias na área.

O presente artigo é apresentado como segue. Na seção 2 são discutidos os estudos relacionados. A seção 3 é composta pela fundamentação teórica. Na seção 4, é descrito o processo de revisão, a seção 5, discute os resultados encontrados. E, por fim, na seção 6, são apresentadas as considerações finais.

2. Trabalhos Relacionados

Gerenciamento de riscos durante o desenvolvimento de projetos de software não é um assunto recente. A prática foi inicialmente avaliada por Boehm (1988), que publicou um estudo mostrando que a aplicação de gerenciamento de riscos é capaz de contribuir de maneira positiva para o sucesso de projetos de *software*.

O trabalho realizado por Andrade (2010) incluiu um estudo sistemático exploratório sobre abordagens, práticas e estratégias de gerenciamento de riscos em desenvolvimento de *software*. O estudo aborda uma série de publicações relevantes que discutem o gerenciamento de riscos no desenvolvimento de software, embora a pesquisa não se limite ao Brasil.

Santos (2011) também utiliza a engenharia de *software* empírica para realizar seu estudo. Através de uma revisão sistemática, foi possível reunir ferramentas utilizadas para o gerenciamento de riscos em desenvolvimento de *software*.

Lobato (2012) discutiu o tema através de uma série de trabalhos empíricos focados em gerenciamento de riscos: (i) duas revisões sistemáticas; (ii) uma síntese narrativa relacionada às revisões sistemáticas; (iii) dois estudos de caso, sendo um na academia e outro na indústria; (iv) um estudo *cross case* realizando uma comparação entre os estudos de caso desenvolvidos; (v) um *survey* com especialistas, relacionando a academia e a indústria; e, por fim, (vi) uma síntese de evidências que teve como objetivo comparar os resultados dos estudos realizados a fim de se desenvolver uma abordagem para gerenciamento de riscos.

Tais trabalhos apresentam estudos empíricos aplicados ao gerenciamento de riscos a fim de reunir informações sobre abordagens, ferramentas, e estratégias de gerenciamento de riscos, sendo assim considerados importantes para o trabalho desenvolvido, visto que apresentaram resultados relevantes em relação ao contexto de gerenciamento de riscos durante o desenvolvimento de software.

3. Fundamentação Teórica

O termo *risco* se refere a qualquer evento que gere consequências negativas a um projeto [Boehm 1991], podendo ser classificado como um evento ou condição incerta que, se ocorrer, terá um efeito negativo sobre pelo menos um objetivo do trabalho, como relatado no *Project Management Body of Knowledge* – (PMI, 2004).

Para avaliar o gerenciamento, a abordagem desenvolvida por Boehm (1989) foi considerada neste trabalho. Para tal consideração, foi avaliado o número de autores da área que utilizam tal trabalho como referência para seus estudos. O estudo em questão apresentou uma abordagem para gerenciar riscos dividida em duas etapas: a etapa de Avaliação de Riscos (composta por identificação, análise e priorização de riscos), e a

etapa de Controle dos Riscos (composta pelo plano de gerenciamento de riscos, a resolução dos riscos e o monitoramento dos riscos).

O processo estratégico deve envolver diferentes *stakeholders* e incluir as tarefas de identificação, análise, planejamento, monitoramento, controle e comunicação de riscos. Os princípios básicos da gerência de risco são os fatores norteadores das atividades que precisam ser desenvolvidas [Gusmão e Moura 2004].

4. Processo de revisão

O processo de revisão incluiu uma busca na forma de revisão sistemática. Eventos científicos realizados no país, bases de dados de bibliotecas digitais, e a internet foram as fontes utilizadas a fim de obter os estudos avaliados.

4.1. Tipos de Busca e Seleção

Para obter resultados consistentes, foram utilizados dois tipos de pesquisa. A primeira de forma manual, incluindo anais de eventos e a internet, enquanto que a segunda consistiu de busca automática englobando repositórios *online* de artigos científicos.

A busca manual em eventos científicos foi realizada em 3 etapas. Na primeira etapa, foi feita uma pesquisa a respeito dos eventos científicos relevantes de Engenharia de *Software* realizados no Brasil. Essa tarefa gerou uma lista com 24 eventos descritos na tabela 1.

Tabela 1. Lista de Eventos Científicos sobre Engenharia de *Software* realizados no Brasil

CBSOFT	CISBE	CLEI	ESELAW
IDEAS	IHC	LA-WASP	SBCARS
SBCCI	SBES	SBLP	SBQS
SBSC	SBSEG	SBSI	SEAS
SEMISH	WDBC	WER	WGRS
WoLLIC	WPERFORMANCE	WSCAD	WTF

Na segunda etapa foi realizada uma busca manual nos anais dos eventos selecionados, restringindo-se ao período de 2002 a 2012. Artigos e tutoriais publicados foram considerados nesta pesquisa. É importante destacar que alguns destes eventos não ocorrem anualmente.

A terceira etapa foi realizada a fim de restringir a busca ao objetivo do trabalho. O primeiro critério de seleção baseou-se na leitura dos títulos das publicações resultando em uma lista com 36 artigos que mencionavam riscos, gerenciamento de riscos ou ameaças em desenvolvimento de *software*. Para tanto, a leitura dos resumos das publicações selecionadas foram avaliadas e, assim, foi possível obter uma nova lista de resultados contendo 21 publicações. A partir desta fase, critérios de inclusão e exclusão foram aplicados às publicações e, por meio de leitura completa das publicações a busca manual resultou em 8 publicações.

Como critérios de inclusão foram consideradas as pesquisas que aplicam o gerenciamento de riscos em projetos de *software*, e as abordagens e ferramentas destinadas à atividade de gerenciamento de riscos em *software*. Como critérios de

exclusão foram retidos pesquisas que apenas referenciavam a atividade de gerenciamento de riscos além de trabalhos com informações inconsistentes e que não apresentavam metodologias para gerenciar os riscos em projetos de *software*.

A busca automática teve como princípio uma pesquisa similar à busca manual para inclusão e exclusão de resultados porém, o método de pesquisa foi definido de forma diferente. Inicialmente, as *strings* de busca foram definidas através de reuniões com *experts* em gerenciamento de riscos, sendo estas combinadas com expressões booleanas *AND* e *OR* a fim de obter expressões que definam o escopo da pesquisa. A tabela 2 representa a lista de *strings* utilizadas na pesquisa.

Tabela 2. Strings de busca

Strings de busca na Língua Portuguesa	Strings de busca na Língua Inglesa
Brasil AND (“gestão” OR “gerenciamento” OR “gerência”) AND “riscos” AND “software”	Brazil AND “management” AND “risks” AND “Software”
Brasil AND (“gestão” OR “gerenciamento” OR “gerência”) AND “riscos” AND “projetos de software”	Brazil AND “management” AND “risks” AND “software projects”
Ferramentas AND Brasil AND (“gestão” OR “gerenciamento” OR “gerência”) AND “riscos” AND “projetos de software”	Brazil AND Tools AND “management” AND “risks” AND “software projects”
Brasil AND Abordagens AND (“gestão” OR “gerenciamento” OR “gerência”) AND “riscos” AND “projetos de software”	Brazil AND Approaches AND “management” AND “risks” AND “software projects”
Brasil AND Modelos AND (“gestão” OR “gerenciamento” OR “gerência”) AND “riscos” AND “projetos de software”	Brazil AND Models AND “management” AND “risks” AND “software projects”
Brasil AND (“gestão” OR “gerenciamento” OR “gerência”) AND “incertezas” AND “projetos de software”	Brazil AND (“management” OR “management” OR “management”) AND “uncertainty” AND “software projects”

As máquinas de busca selecionadas são compostas por campos de busca avançados, possibilitando a busca automatizada e, seleção de filtros de resultados. Assim, para esta pesquisa foi aplicado um filtro de classificação focado na relevância das publicações. No contexto deste estudo, as ferramentas utilizadas foram o Google Acadêmico, Scielo Brasil, ACM e IEEE *Xplore*.

Na terceira etapa foi encontrado um valor médio de 7205 resultados por *strings* considerando a repetição de resultados para cada string. Apenas os estudos que continham título resultante do casamento de padrão das strings foram utilizados para a pesquisa. Devido à aplicação do filtro de relevância das máquinas de busca, optou-se por avaliar os 100 resultados com maior relevância de acordo com cada *string*. Para realizar a seleção dos estudos, foi necessário avaliar os resumos de todos os trabalhos a fim de adequar aos critérios de inclusão e exclusão apresentados anteriormente. Para os trabalhos selecionados, realizou-se uma leitura do resumo de cada publicação e por sequência, publicações que se enquadraram no contexto do estudo receberam atenção especial através de leitura completa.

4.2. Avaliação da Qualidade

A avaliação da qualidade aplicada ao trabalho foi realizada com base em um esquema de pontuação. Para pontuar as perguntas, utilizou-se a seguinte estratégia: respostas que satisfazem plenamente as perguntas receberam o valor 1; respostas que satisfazem parcialmente as perguntas receberam o valor 0,5 e, por fim, respostas que não satisfazem as perguntas recebem o valor 0.

A figura 1 representa as questões propostas para avaliar a qualidade dos estudos selecionados.

ID	Questões de Critério de Qualidade	Pontuações
QC1.	O estudo aborda gerenciamento de riscos em projetos de software?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC2.	O estudo apresenta as principais características de gerenciamento de riscos?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC3.	O estudo apresenta fundamentação quanto a aplicação de gerenciamento de riscos em projetos de software?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC4.	O estudo apresenta resultados quanto à aplicação de gerenciamento de riscos?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC5.	O estudo classifica os riscos?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC6.	O estudo apresenta características de contingência e mitigação dos riscos?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0
QC7.	O estudo é aprofundado?	Sim = 1.0 / Em Parte = 0.5 / Não = 0.0

Figura 1. Questões de critério de qualidade

Com a aplicação da revisão sistemática foi possível obter um total de 37 pesquisas relacionadas ao tema deste estudo. Dentre as 37 pesquisas foram destacados os trabalhos que apresentaram maior relevância e maior incidência entre os meios de busca.

5. Resultados

O estudo empírico aplicado neste trabalho mostrou resultados importantes. Dentre as pesquisas encontradas, algumas foram consideradas mais relevantes devido à sua classificação por meio de máquinas de busca e, também, com base no índice de incidência apresentados nas buscas realizadas.

A seguir são apresentadas em ordem cronológica de tempo as pesquisas que obtiveram maior destaque neste estudo:

- Machado (2002) apresentou um método de gerência de risco, chamado de *A-Risk*. O *A-Risk* é focado na identificação e quantificação de riscos de prazo de projeto, que pode ser aplicado antes e durante o desenvolvimento do projeto, ou seja, em todas as suas fases.
- Costa *et. al.* (2004) desenvolveu uma abordagem para avaliar projetos com riscos de nível global por meio de um estudo empírico planejado e executado para quantificar a importância relativa dos riscos do projeto para uma categoria específica do sistema.
- Fontoura *et. al.* (2004) propôs um paradigma para gerir riscos em projetos de *software* conhecido com GQM (*Goal/Question/Metric*). Tal paradigma é baseado

em métricas do processo de *software* utilizadas para acompanhar o progresso dos fatores de risco, possibilitando ao gerente de projeto tomar ações corretivas, quando necessário e no momento adequado.

- Gusmão e Moura (2004) apresentaram uma análise detalhada sobre as atividades de gerenciamento de riscos presente nos modelos de qualidade de *software*. O trabalho demonstra que não há um processo padrão para gerenciar riscos, apenas práticas consolidadas, além de não existir indicação de processos e métodos que podem ser utilizados na execução de atividades de gerenciamento, fazendo com que as organizações ajustem seus processos à atividades de gerenciamento de risco, de acordo com sua realidade.
- Leopoldino e Bezerra (2004) desenvolveram um estudo sobre avaliação de riscos em projetos de *software* julgando riscos encontrados no desenvolvimento de *software* por meio de experiências de desenvolvedores e gerentes de projetos.
- Oliveira e Bastos (2005), apresentaram um método para aplicação de gerência de riscos em projetos de *software* com o objetivo de propor um processo de gerência de risco aplicado às dimensões de custo, prazo e qualidade de projetos desenvolvidos para integração com um processo de desenvolvimento de *software* focado em sistemas de informação.
- Knob *et. al.* (2006), desenvolveu a ferramenta *RiskFree* com o objetivo de auxiliar equipes de projeto de desenvolvimento de *software*;
- Leme (2007) desenvolveu uma estratégia e um protótipo de ferramenta de apoio ao gerenciamento de riscos para um ambiente de desenvolvimento distribuído de *software* integrando ao ambiente DiSEN.
- Mendes *et. al.* (2007) identificou na literatura os problemas e riscos mais citados com o intuito de gerar uma lista dos dez riscos mais recorrentes além de propor um método para minimizar o impacto e/ou probabilidade de ocorrência do risco.
- Dantas *et. al.* (2009) descreve um módulo de controle de riscos, conhecido como Risys, este módulo de gerenciamento de riscos é aplicado na execução de projetos desenvolvidos pela utilização do *Webbased Groupware on Project Management System* (WGPMS).
- Falbo (2010) apresenta uma ontologia de riscos de *software*, estabelecendo uma conceituação comum acerca do domínio de riscos de *software*, a qual pode ser adotada por organizações de *software* como um vocabulário básico para se falar sobre riscos.
- Lima *et. al.* (2010) propuseram modificações na ferramenta *Risys* de modo a ampliar o tratamento de riscos no WGPMS.
- Lobato *et. al.* (2010) apresenta uma revisão sistemática sobre gerenciamento de riscos onde são apresentadas algumas estratégias e passos que podem ser utilizados durante o gerenciamento de riscos para desenvolvimento de *software*.

Com base nas pesquisas apresentadas, é possível afirmar que a área de gerenciamento de risco necessita de mais atenção relacionado a qualidade de pesquisas disponíveis na literatura. Os principais problemas identificados foram questões sobre as atividades de gerenciamento de riscos que nem sempre foram tratadas e etapas para

realizar o gerenciamento de riscos que não foram bem descritas. Estas desvantagens impactam nos resultados obtidos neste trabalho, uma vez que o estudo obteve um índice alto de respostas negativas em comparação com as respostas positivas.

6. Considerações Finais

O principal objetivo deste trabalho foi investigar o estado da arte em pesquisas sobre gerenciamento de riscos através de um estudo de revisão sistemática. O estudo desenvolvido foi o primeiro a apresentar uma revisão sistemática com o foco em pesquisas desenvolvidas utilizando gerenciamento de riscos em projetos de *software* no Brasil. Este trabalho serve como referência a trabalhos futuros, uma vez que apresenta diretrizes de pesquisa quanto às atividades de gerenciamento de riscos e às etapas para controlar os riscos.

Este o estudo foi capaz de fornecer medidas e ideias úteis sobre o gerenciamento de riscos tanto na indústria quanto na academia além de vantagens e técnicas para realizá-lo. Ainda assim, é importante destacar que os recursos utilizados nas pesquisas não foram bem explorados apesar do número crescente de estudos sobre este tema nos últimos anos, indicando que o gerenciamento de riscos ainda se trata de um alvo de pesquisa na engenharia de software.

Como ameaça a validade da pesquisa, pode-se citar que alguns trabalhos relevantes podem não ter sido selecionados, uma vez que as máquinas de busca não indexam artigos de todas as conferências brasileiras.

Para trabalho futuro, pretende-se investigar profundamente os recursos utilizados pelas pesquisas a fim de avaliar e propor melhorias para abordagens e ferramentas sobre gerenciamento de riscos em desenvolvimento de software.

Referências

- Andrade, H. S. (2010) “*Systematic scoping study on risk management in software development.*” In: Universidade Federal de Goiás, Campus Catalão – UFG, Catalão, Brasil.
- Boehm B. W. (1988) “*A spiral model of software development and enhancement*” In: IEEE Computer, 21(5), 61-72.
- Boehm B. W. (1989) “*Software risk management*” Em: IEEE Computer Society Press.
- Boehm B. W. (1991) “*Software risk management: Principles and practices*” Em: IEEE Software.
- Costa H. R., Barros M. de O. e Travassos G. H. (2004) “*Evaluating risk factors in software projects*” Em: Eselaw – Experimental Software Engineer Latin American Workshop.
- Dantas B. T., David J. M. N., Avelar A. J. N., Ferreira L. A. S. e Jesus L. M. S. (2009) “*Risys – Uma ferramenta de apoio à gerência de riscos em um ambiente colaborativo de gestão de projetos*” Em: SBSC – Simpósio Brasileiro de Sistemas Colaborativos.
- Falbo R. A. (2010) “*Uma ontologia de riscos de software*” In: SBQS – Simpósio Brasileiro de Qualidade de Software.

- Fontoura L. M., Price R. T. e Phil D. (2004) “Usando GQM para gerenciar riscos em projetos de *software*” Em: SBES – XVIII Simpósio Brasileiro de Engenharia de *Software*.
- Gusmão, C. M. G. and Moura, H. P. (2004) “Gerência de risco em processos de qualidade de *software*: uma Análise Comparativa” Em: Anais do III Simpósio Brasileiro de Qualidade de *Software*, Brasília – DF – Brasil.
- Knob F., Silveira F., Orth A. I., Prikladnicki R. (2006) “*RiskFree* – Uma ferramenta de gerenciamento de riscos baseada no PMBOK e aderente ao CMMI” Em: SBQS - Simpósio Brasileiro de Qualidade de *Software*.
- Leme L. H. R. (2007). “Uma estratégia para apoiar o gerenciamento de riscos em um ambiente distribuído de desenvolvimento de *software*”, <http://200.17.137.109:8081/xiscanoe/Members/vollare/risk-analysis-and-productivity-in-software-engineering/>. Acesso em 01/02/2013.
- Leopoldino C. B. e Bezerra C. (2004) “Avaliação de riscos em desenvolvimento de *software*” Em: Dissertação de Mestrado, Programa de Pós-Graduação em Administração – Universidade Federal do Rio Grande do Sul – UFRGS.
- Lima M. P., David J. M. N., Dantas B. T. (2010) “Gerenciamento de riscos e contexto em ambientes colaborativos de gestão de projetos de desenvolvimento de *software*” Em: SBSC – Simpósio Brasileiro de Sistemas Colaborativos.
- Lobato L. L., O’Leary, P., Almeida, E. S., Meira, S. R. L. (2010) “*The Importance of Documentation, Design and Reuse in Risk Management for SPL*” Em: 28th ACM International Conference on Design of Communication (SIGDOC), São Carlos.
- Lobato L. L. (2012) “*An Approach for Risk Management in Software Product Line*” Tese de doutorado. Em: Universidade Federal de Pernambuco, PE, Brasil.
- Machado C. A. F. (2002). “*A-RISK* : Um método para identificar e quantificar riscos de prazo em projetos de desenvolvimento de *software*”, <http://www.inforede.net/Technical/Business/IT/PM%20Risco%20em%20Desenvolvimento.pdf>. Acesso em 28/01/2013.
- Mendes F. F., Oliveira J. L., Fernandes P. G. e Souza A. S. (2007) “Análise de riscos na implantação de melhorias de processos de *software*”, [http://www.softex.br/portal/softexweb/uploadDocuments/_mpsbr/\[03\]%20Anali seRiscosImplantacao_FINAL.pdf](http://www.softex.br/portal/softexweb/uploadDocuments/_mpsbr/[03]%20Anali seRiscosImplantacao_FINAL.pdf), acesso em 01/02/2013.
- Oliveira G. da C. e Bastos R. (2005) “No-Risk - Um Método para Aplicação de Gerência de Riscos em Projetos de *Software*” Em: SBSI – Simpósio Brasileiro de Sistemas de Informação, Florianópolis, 2005.
- PMI – *Project Management Institute* (2004) “*A guide to the project management body of knowledge*” 3ª edição, PMBOK Guide.
- Pressman, R. S. (1995) “Engenharia de *software*” São Paulo: Ed. Makron Books.
- Santos, M. R. (2011) “Um estudo sistemático sobre gerenciamento de riscos para projetos de *software*” Trabalho de conclusão de curso. Em: Universidade Federal de Goiás - UFG, Catalão, Brasil.

Concepção de uma Ferramenta Parser para Extração de dados do metamodelo ECore

Ítalo de Pontes Oliveira¹, Lucas Araújo Ramos¹, Rhavy Maia Guedes¹

¹ Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) – Campina Grande, Paraíba – Brasil

italo.oliveira@ieee.org, {lucas.du.a7x@gmail.com, rhavy.maia}@gmail.com

Abstract. *The Software Engineering offers clear and objective techniques for systems building. There are model based approaches for the software development cycle, such as the MDA. The OMG defines models as graphs that provide important information about the system. It is necessary that modeling teams seeking qualify and quantify the entities that are involved in the process. For this, they need to extract information from files based on the standard MOF metamodels. Thus, this study aims to perform data extraction for ECore metamodel by creating a tool-based software that allows the analysis of the elements contained in the model.*

Resumo. *A Engenharia de Software oferece técnicas claras e objetivas na construção de sistemas. Existem abordagens baseadas em modelo para o desenvolvimento do ciclo de software como, por exemplo, o MDA. A OMG define os modelos como representações gráficas que trazem informações importantes sobre o sistema. É necessário que as equipes de modelagem busquem qualificar e quantificar as entidades que estão envolvidas no processo. Para isso, necessitam extrair informações de arquivos de metamodelos baseados no padrão MOF. Desse modo, o presente trabalho tem como objetivo realizar a extração de dados do metamodelo ECore através da criação de uma ferramenta baseada em software que permitirá a análise dos elementos contido no modelo.*

1. Introdução

A engenharia de software oferece técnicas claras e objetivas na construção de sistemas, a fim de obter melhores resultados. Para isso, o modelo de processo de software é composto por atividades de apoio como, por exemplo, o Levantamento e Análise de Requisitos, a Modelagem do Sistema, a Implementação e os Testes [Presman 2010]. A Análise de Requisitos é o processo de avaliação das funcionalidades do sistema, e tem como objetivo evitar redundância e esclarecer os procedimentos a serem realizados pelo software. A Modelagem do Sistema constrói uma visualização prévia das funcionalidades de abrangência. O Desenvolvimento do Sistema define a estrutura de código. Os Testes tratam as falhas e exceções do desenvolvimento.

Para Qiuyan (2012), a modelagem permite uma representação mais clara dos requisitos a serem implementados no sistema. Além disso, serve como documentação de

contrato entre a equipe de desenvolvimento e o cliente. A Object Management Group (OMG) [OMG 2013] define os modelos como representações gráficas que trazem consigo informações importantes sobre o sistema. Segundo Liu (2005), os modelos são estruturados a partir de seus metamodelos que contém uma sintaxe abstrata e semântica estática de uma linguagem de modelagem específica de domínio.

Além disso, Piers (2012) ressalta a possibilidade de realizar transformações automáticas em modelos com o objetivo de melhorar a qualidade, evitar erros e economizar esforços no processo de desenvolvimento de software. A técnica de transformação automatizada permite que através de um modelo de entrada seja gerado um ou mais modelos de saída. O procedimento possibilita que as alterações sofridas em processos organizacionais modelados sejam rapidamente refletidas no software. Para isso, a OMG define padrões e metodologias para o processo de transformações de modelos como, por exemplo, o *Meta Object Facility* (MOF) que define uma arquitetura para criação de metamodelos, e o *Atlas Transformation Language* (ATL) linguagem procedural capaz de executar transformações unidirecionais.

No entanto, antes da execução da transformação do modelo ou modelagem, as equipes buscam qualificar e quantificar as entidades que estão envolvidas no processo, a fim de identificar relações e dependências dos elementos e o impacto causado pelas mudanças. A análise geralmente é realizada através das imagens dos diagramas gerados pelas ferramentas. No entanto, a análise visual é passível de erro, pois os diagramas podem estar representados em um nível de abstração alto.

Para extrair as informações contidas em um metamodelo é necessário que seja interpretados a estrutura de metadados do arquivo baseado em XML *Metadata Interchange* (XMI). O XMI permite o intercâmbio de metadados entre ferramentas baseadas no padrão da OMG como, por exemplo o *Eclipse Model Framework* (EMF) que utiliza o metamodelo ECore.

Neste sentido, o presente trabalho tem como objetivo realizar a extração de dados do metamodelo ECore através da criação de uma ferramenta baseada em software que permitirá a análise dos elementos contidos no modelo. Essas informações darão subsídios para que equipes de modelagem tomem decisões acerca da transformação de modelos.

2. Background

O *Model Driven Engineering* (MDE) é uma tendência da engenharia de software que reúne todas as abordagens baseadas em modelo para o ciclo de desenvolvimento de software. [Rensink 2006]. O *Model Driven Architecture* (MDA) faz parte do conjunto proposto pela MDE. O Objetivo do MDA é definir a estrutura e o comportamento do sistema em um nível de abstração mais alto que desconsidera tecnologias e implementações subjacentes. Sua proposta principal é diminuir o contato com código-fonte, deixando o esforço voltado para a análise e desenvolvimento de modelos. Dessa forma, o modelo deixa de ser apenas um guia, uma referência para o desenvolvimento, e passa a compor o software assim como o código-fonte [Moraes 2002].

Para definir sua estrutura semântica e sintática, o modelo pode utilizar um ou mais metamodelos. Segundo a OMG, o metamodelo confere vantagens na implementação e interoperabilidade de modelos, pois define regras, restrições e teorias úteis para a modelagem.

2.1. Estrutura de um modelo

A OMG padronizou a arquitetura de metamodelo através da criação do *Meta Object Facility* (MOF), desenhada quatro meta-camadas. O MOF é capaz de estruturar sintaticamente e semanticamente um conjunto de modelos [Silva 2008]. A Figura 1 mostra as camadas que compõe o padrão MOF, é possível identifica-las de baixo para cima pelos índices M0 até o M3.

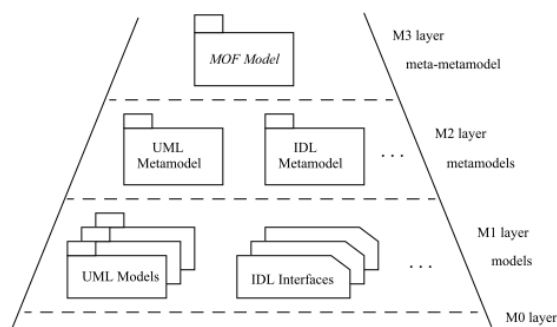


Figura 1. Arquitetura do Metamodelo MOF [OMG 2013].

O nível M0 é a camada mais baixa da arquitetura que constituem as instâncias concretas dos elementos representados num modelo; O nível M1 de Modelos responsável por representar os elementos que constituem as abstrações definidas pelo utilizador da UML no seu processo de modelagem; O nível M2 de Metamodelo é composto por elementos que constituem as abstrações definidas pelos criadores da UML, ou seja, estruturas para definição dos modelos (Por exemplo: Classe, Atributo, Método, Relacionamento, etc.); Por fim, o nível M3 de Meta-metamodelo que permite definir os elementos que serão responsáveis pela especificação dos metamodelos (Por exemplo: MetaClasse, MetaAtributo, MetaMétodo, MetaRelacionamento, etc.) [Silva 2008][Rensink 2006]. Durante a implementação do padrão MOF as ferramentas podem optar em reduzir ou aumentar a quantidade de camadas definidas pela OMF.

Baseado nas especificações da OMG o EMF permite a construção de ferramentas e aplicações baseadas em modelo de classe baseado na UML [Steinberg et al 2009]. O EMF converte os modelos criados para o formato ECore que contém informações acerca das classes definidas pelo modelo [Vogel, 2008]. É importante ressaltar que o metamodelo definido pelo ECore é baseado no padrão *Meta-Object Facilities* (MOF). Essa característica permite maior integração do EMF com outras ferramentas que também utilizam o padrão MOF.

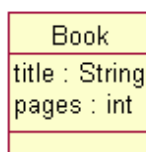


Figura 2. Modelo de classe gerado pelo EMF [EMF 2013].

É possível observar na A Figura 2 um simples modelo de classe gerado através do EMF. A classe denominada *Book* contém o atributo “*title*” do tipo *String*, e “*pages*” do tipo *int*. O modelo da Figura 2 é descrito através da meta-informações presentes no documento do XML *Metadata Interchange* visto na Figura 3.

```

<xsd:schema targetNamespace="http://library.ecore"
  xmlns="http://library.ecore" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
  <xsd:complexType name="Book">
    <xsd:sequence>
      <xsd:element name="title" type="xsd:string"/>
      <xsd:element name="pages" type="xsd:integer"/>
    </xsd:sequence>
  </xsd:complexType>
</xsd:schema>
  
```

Figura 3. Representação textual XML Schema para um modelo EMF [EMF 2013].

A Figura 4 representa as quatro entidades básicas do metamodelo ECore. O ECore é composto pela entidade EClass, EAttribute, EDataType, EReference. A EClass modela uma classe que pode possuir um identificador, cardinalidade e referências. A EAttribute modela um atributo que pode possuir nome e tipo. A EReference, representa a extremidade de uma associação entre as classes. A EDataType modela um tipo de atributo, podendo ser um tipo de dado primitivo ou composto.

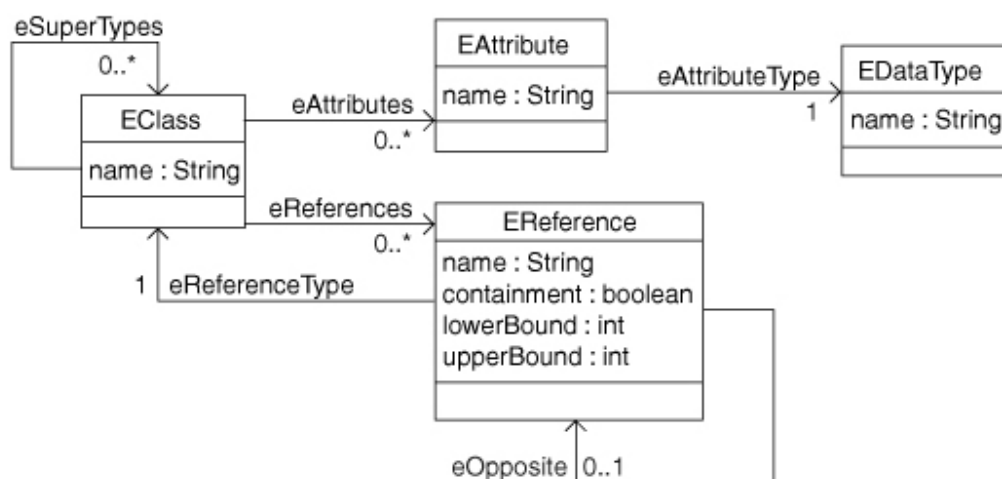


Figura 4. Diagrama de Classes para a representação feita por um modelo ECore.

3. Materiais e Métodos

O trabalho foi desenvolvido por um grupo de Iniciação Científica que investiga temáticas relacionadas a Modelagem Dirigida a Modelos (MDD) no Instituto Federal de Educação,

Ciência e Tecnologia da Paraíba campus Campina Grande. O grupo é formado por estudantes do curso Superior em Telemática e pelo professor orientador. Durante o período de revisão da literatura foram levantados aspectos importantes sobre a transformação automática de modelos como, por exemplo, a análise estática e extração de dados de metamodelos.

3.1 Ferramentas de modelagem

A pesquisa utilizou o ambiente de modelagem baseado no (EMF) para desenvolver o modelo presente na Figura 5. O modelo representa o Sistema de Orientação Docente-Discente que servirá como artefato de análise da ferramenta de análise estática e extração de informações de modelos.

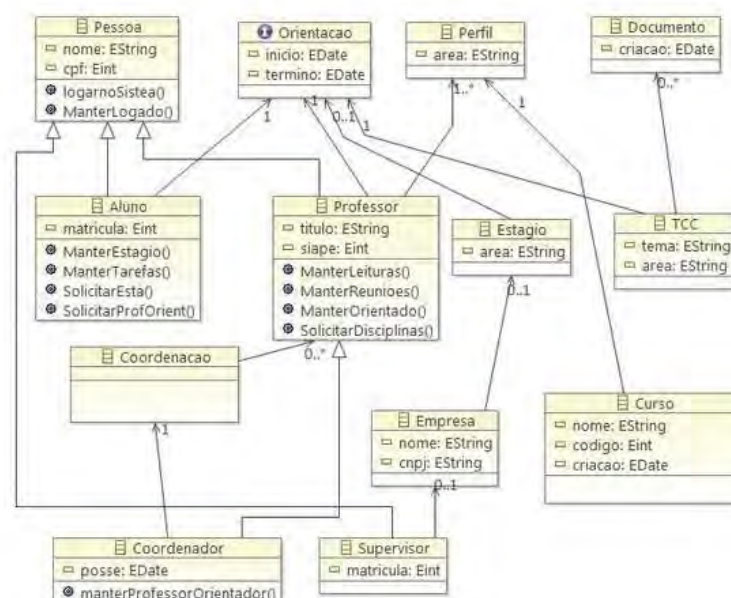


Figura 5. Modelo Sistema de Orientação Docente-Discente.

3.2. Especificação de Requisitos

Tabela 1 - Lista de requisitos para construção do sistema

Requisito	Descrição
Extrair dados do ECore	Extrair as informações contidas no arquivo XMI que representa o modelo de classe baseado em ECore. A extração resultará na criação de uma instância de objeto Java composto por uma lista de Classes que seguirá a definição do modelo EClass.
Exibir Relacionamento	Exibir informações sobre as classes que compõe o modelo. Será possível visualizar os atributos, agregações e

	heranças presentes nos relacionamentos bem como sua cardinalidade.
Gerar Relatório para Análise	Gerar um relatório capaz de quantificar e analisar as características do modelo. O documento é sintetizado e de fácil interpretação, pois auxiliará no processo de transformação de modelo.

3.3. Implementação das funcionalidades

O sistema foi estruturado a partir da arquitetura prevista pelo Model-View-Controller (MVC) e codificado através da linguagem de programação Java. A codificação dos requisitos de Extração de dados do ECore e Geração de relatório realizou-se no ambiente de desenvolvimento do Eclipse. No entanto, para a construção das interfaces gráficas baseadas em Swing utilizou-se a ferramenta NetBeans.

Como forma de modularizar os comportamento referentes a Extração de dados do ECore a equipe resolveu criar uma API base responsável por transformar o arquivo XMI, baseado no ECore, em instância de objeto que representa o diagrama de classe mostrado na Figura 4.

4. Resultados

A aplicação inicia o procedimento de interpretação da estrutura do modelo de classe através da leitura de um arquivo XMI disponibilizado pela equipe de modelagem. A interpretação desde arquivo gera a instância de objeto que contem a lista de classe com suas características e comportamento. É possível observar na Figura 6 a exibição de todas as entidades existentes no modelo da Figura 5. A interface permite que a equipe de modelagem acesse as opções de atributos, métodos, relacionamentos inerente a cada classe. A opção Gerar Relatórios fornece um detalhamento quantitativo de relacionamentos, heranças, quantidades de atributos e comportamentos presentes em todo o modelo.

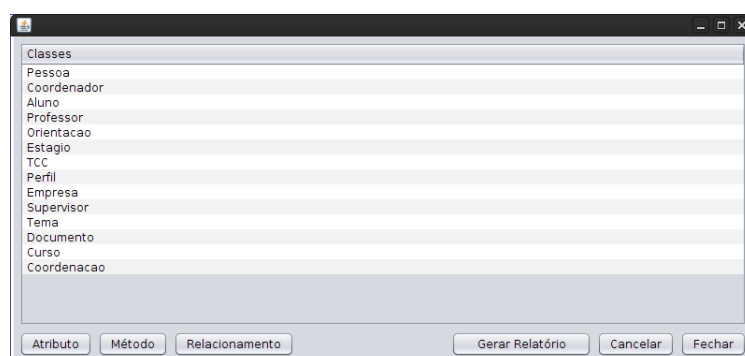


Figura 6. Tela exibindo as classes de um modelo de entrada.

5. Considerações Finais

A ferramenta poderá auxiliar empresas de desenvolvimento ou pesquisadores durante a realização da avaliação de estrutura de modelos baseados em ECore. Será possível que

uma equipe de engenharia de software avalie, através da ferramenta desenvolvida, a maturação de um modelo em linha de desenvolvimento de software, evitando assim, erros durante o processo de codificação de sistemas humanos.

Referências

- Piers, W. and Bruneliere, H. (2012). ATL Concepts. Disponível em: <<http://wiki.eclipse.org/ATL/Concepts>>. Acessado em: 10 fev. 2013.
- Vogel, L. (2008). Eclipse EMF. Disponível em: <<http://www.vogella.com/articles/EclipseEMF/article.html>>. Acessado em: 19 fev. 2013.
- Burbeck, S. (1987). Applications programming in smalltalk-80(tm): How to use model-view-controller (mvc).
- Castro, S. (2008/2009). Modelos de código java eclipse java model e emf java model. Technical report, Instituto Superior de Engenharia do Porto.
- EMF (2013). Emf. the eclipse modeling framework (emf) overview. Disponível em: <http://help.eclipse.org/indigo/nav/220>. Acessado em: 07 jan. 2013.
- Félix, R. (2010). Dsltranslator - ferramenta para transformação de modelos. Master's thesis, Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Departamento de Informática.
- Kleppe, A. G., Warmer, J., and Bast, W. (2003). MDA Explained: The Model Driven Architecture: Practice and Promise. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.
- Kreutz, M. and Silva, D. R. C. (2013). Metr pole digital. Disponível em: http://www.metroledigital.ufrn.br/aulas_avancado/web/disciplinas/desktop/aula03.html. Acesso em: 18 jan. 2013.
- Lapenda, R., Madruga, P., and Loniewski, G. (2008). Utilizando transformação de modelos para o desenvolvimento de softwares de qualidade.
- Liu, J., He, K., Li, B., He, C., and Liang, P. (2005). A transformation definition metamodel for model transformation. In Information Technology: Coding and Computing, 2005. ITCC 2005. International Conference on, volume 2, pages 373 – 378 Vol. 2.
- Lucr dio, D. (2009). Uma Abordagem Orientada a Modelos para Reutiliza  o de Software. PhD thesis, Instituto de Ci ncias Matem ticas e de Computa  o. Universidade de S o Paulo.
- Moraes, T. (2012). Extra  o do comportamento especificado em modelos uml usando o eclipse modeling framework.
- Oliveira,  . P., Bezerra, T. S., Ramos, L. A., Guedes, R. M. (2012). Uso de modelos de integra  o para transforma  es de modelos baseados em metamodelos. VII CONNEPI.
- OMG (2013). Documents associated with mof version 2.0. Disponível em: <http://www.omg.org/spec/MOF/2.0/>. Acesso em: 07 jan. 2013.

- Pressman, R. (2010). *Software engineering: a practitioner's approach*. McGraw-Hill higher education. McGraw-Hill Higher Education.
- Qiuyan, L., Jie, T., Qihong, P., Ji, W., and Chao, L. (2011). Automatic transformation technology from aadl model to uml model. In *Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on*, pages 255–258.
- Rahimi, S. (2010). Specification of uml model transformations. In *Software Testing, Verification and Validation (ICST), 2010 Third International Conference on*, pages 323–326.
- Selic, B. (2003). The pragmatics of model-driven development. *IEEE Softw.*, 20(5):19–25.
- Silva, A. d., Videira, C., and Atlântico, C. (2008). *UML, metodologias e ferramentas CASE: linguagem de modelação UML, metodologias e ferramentas CASE na concepção e desenvolvimento de sistemas de informação*. Number v. 1. Centro Atlântico.
- Sommerville, I. (2003). *Engenharia de Software*. Addison Wesley, São Paulo, SP, 6 edition. Tradução André Maurício de Andrade Ribeiro; Revisão técnica Kechi Hiramã.
- Steinberg, D., Budinsky, F., Paternostro, M., and Merks, E. (2009). *EMF: Eclipse Modeling Framework 2.0*. Addison-Wesley Professional, 2nd edition.
- Zhang, J., Liu, S., Wang, X., and Qin, T. (2008). A model transformation classification method used in qvt. In *IT in Medicine and Education, 2008. ITME 2008. IEEE International Symposium on*, pages 464 –468.

Aplicando Padrões de Seleção no Ensino de Programação de Computadores para Estudantes do Primeiro Ano do Ensino Médio Integrado

Alexis V. de A. Leal¹, Deller J. Ferreira²

¹Instituto Federal de Mato Grosso – *campus* Barra do Garças (IFMT)
Estrada de acesso a BR-158, S/N – 78.600-000
Barra do Garças – MT – Brazil

²Instituto de Informática – Universidade Federal de Goiás (UFG)
Caixa Postal 131 – 79-001-970 – Goiânia – GO – Brazil
{alexisleal,deller}@inf.ufg.br

Abstract. *This paper describes an experience related to the use of programming patterns in the teaching of algorithms and computer programming to students the first year of high school in integrated mode. We present and discuss the preliminary results of ongoing research, which addresses the combined use of algorithmic programming patterns with concrete materials and games.*

Resumo. *O presente artigo descreve uma experiência relacionada ao uso de padrões de programação no ensino de algoritmos e programação de computadores, para estudantes do primeiro ano do Ensino Médio na modalidade Integrado. São apresentados e discutidos os resultados preliminares da pesquisa em andamento, que aborda a combinação do uso de padrões de programação de programação e jogos com materiais concretos.*

1 - Introdução

A preocupação com a falta de mão de obra em tecnologia, especialmente na área de Informática, vem chamando muito a atenção nos dias atuais, tanto por parte do governo, quanto pela iniciativa privada. Conforme o MINISTÉRIO DA EDUCAÇÃO [2010], o Governo Federal investirá R\$ 1,2 bilhão nos Institutos Federais de Educação, Ciência e Tecnologia até o ano de 2014. Parte destes recursos será destinada à formação de profissionais de informática de nível superior e médio técnico, conforme Lei 11.892/2008, a prioridade é a formação na modalidade Integrada ao Ensino Médio. Grande parte dos Institutos Federais no Brasil oferece o curso Técnico em Informática Integrado ao Ensino Médio.

Com relação aos cursos onde a base da formação técnica deve ser em programação, o objetivo inicial é o domínio dos conceitos básicos de algoritmos e programação de computadores, justamente para o prosseguimento dos estudos nos anos subsequentes. Grande parte dos estudantes encontram grandes dificuldades, sejam elas decorrentes de uma formação deficitária no Ensino Fundamental ou o total

desconhecimento do foco do curso no qual se está ingressando, já que muitos são exímios usuários de tecnologias computacionais e aliam a escolha do curso a este fator.

Para suprir a demanda dos profissionais de Tecnologia da Informação, é fundamental que se aprimore a qualidade do processo de ensino e aprendizagem nos cursos da área de computação e informática. Um dos gargalos destes cursos é o alto índice de reprovação e/ou evasão, desde o início do curso, em disciplinas de algoritmos e programação [PEREIRA JÚNIOR *et al*, 2005]. Em geral, os estudantes que ingressam nos cursos técnicos integrados são adolescentes, com idade entre 13 e 15 anos. No caso do IFMT, metade das vagas dos cursos é destinada a alunos que cursaram todos os nove anos do Ensino Fundamental em escolas públicas e a outra metade, para o chamado acesso universal [IFMT, 2011], que são as vagas destinadas a alunos que, ou não cursaram todo o Ensino Fundamental em instituições de ensino públicas ou são egressos de instituições privadas.

Neste sentido, dadas às particularidades dos estudantes em questão, faz-se necessária a busca de novas abordagens e metodologias de ensino, em contraponto ao ensino tradicional de algoritmos e programação de computadores. No chamado modo tradicional, o conteúdo de tal disciplina é apresentado de forma expositiva concentrando-se na solução de problemas. Conforme apresentado em PEREIRA JUNIOR & RAPKIEWICZ [2004], no modo tradicional, o professor apresenta o conteúdo através do chamado *portugol*, apresenta alguns exemplos e propõe exercícios para a turma. Segundo BORGES [2000], o modo tradicional não consegue facilmente motivar os alunos a se interessar pela disciplina, pois não é claro para os alunos a importância de certos conteúdos para sua formação.

Desta forma, este artigo descreve uma primeira experiência, realizada no primeiro bimestre de 2012, entre os meses de abril e junho, numa turma de primeiro ano do Ensino Médio Integrado a Informática. A experiência ocorreu no IFMT *campus* Barra do Garças, na disciplina Algoritmos e Lógica de Programação, utilizando quatro padrões de programação de seleção para resolução dos problemas apresentados em uma lista de exercícios.

O restante deste artigo está organizado da seguinte forma. A Seção 2 faz um estudo do uso de padrões para ensino de programação. O desenvolvimento do trabalho é apresentado na Seção 3. Na Seção 4, apresentamos a conclusão e as possibilidades de trabalhos futuros e por último, na seção 5, agradecemos aos órgãos que apoiaram esta pesquisa.

2. Ensino e Padrões de Programação

Estudantes novatos em programação apresentam muitas dificuldades, tais como problemas relacionados com a capacidade de abstração, falta de habilidades para lidar com problemas e, no caso de estudantes recém-saídos do nono ano do Ensino Fundamental, problemas de aprendizagem, principalmente em conteúdos de Ciências Exatas. Dois grandes desafios para o estudante, conforme WINSLOW [1996] são aprender uma nova linguagem de programação e aprender como criar uma solução computacional que resolva um dado problema. Com relação ao primeiro problema, aprender as regras de sintaxe e semântica da linguagem torna-se relativamente

complexo, ainda mais se considerarmos que os comandos, como no caso da linguagem C++, são em inglês. Entretanto, o maior desafio para o programador iniciante, realmente é interpretar o problema, resolvê-lo algorítmicamente e codificar o algoritmo em uma linguagem de programação.

Programadores experientes conseguem, a partir do problema a ser resolvido, identificar qual estrutura usar, quais os tipos de dados envolvidos, bem como outras formas de resolver o mesmo problema usando estratégias diferentes. Resolução de novos problemas utilizando experiências prévias, adaptando as soluções para resolvê-las, é discutida por JOHNSON e SOLOWAY [1984] através do sistema PROUST, que trabalha com experiências de programação anteriores dos estudantes para o ensino da linguagem de programação Pascal. Estas experiências prévias contêm a base para os padrões de programação, que são soluções que, com frequência, aparecem na resolução de problemas computacionais. Padrões traduzem estratégias de programação criadas por especialistas. Desse modo, os alunos tiram proveito de boas práticas de programação.

Estes padrões elementares, também chamados de padrões pedagógicos de programação [PORTER e CALDER, 2003], contêm o esqueleto no qual o problema será resolvido pelo então iniciante em programação de computadores. Padrões de programação formam uma base inicial de micro soluções que podem ser adaptadas e combinadas pelos alunos durante a elaboração de seus programas.

Um padrão é um modelo que relata uma solução para um problema e fornece informações sobre o contexto em que ele deve ser aplicado [BARROS, *et al*, 2005]. Assim, o estudante iniciando seus estudos em programação, ao se deparar com determinado tipo de problema, pode direcionar sua resolução a partir dos padrões já conhecidos, sendo que o professor tem um papel fundamental neste processo, que será discutido mais adiante. CLANCY e LINN [1999] descreveram como padrões podem ser utilizados no ensino de programação, concluindo que os estudantes necessitam fazer o processo de aprender, criar e reformular padrões, tornando isso uma prática regular em seu processo de aprendizagem de programação. Desta forma, os exercícios devem ser criados pelo professor de forma a estimular o uso de padrões simples, combinações entre diferentes padrões, bem como estimular a criação de novos padrões pelo próprio estudante.

Padrões de programação são soluções para problemas computacionais recorrentes, formando blocos para o desenvolvimento de algoritmos. Padrões de programação se referem à classificação de problemas computacionais de acordo com o seu objetivo. [MULLER, 2005]. Assim, o aluno pode, através de uma classificação identificar, a partir do problema a ser solucionado, qual padrão pode ser utilizado para elaboração do algoritmo que soluciona o problema. Outro problema para o iniciante em programação é a questão da capacidade de abstração. Em HABERMAN e MULLER [2008], é proposto um método de ensino de abstração, combinado com resolução de problemas através de padrões. Neste trabalho, as autoras pesquisaram a influência da combinação de padrões com ensino baseado em tipos abstratos de dados. Sendo assim, o estudo mostrou que é possível combinar ensino baseado em padrões com outros métodos de ensino.

Outros problemas relacionados ao ensino de programação são discutidos por MULLER [2005]. Um deles é a capacidade do estudante de reconhecer similaridades entre problemas, particularmente entre o problema em questão e problemas encontrados e resolvidos previamente. Por conta disso, existe a necessidade de se categorizar, catalogar e descrever os principais padrões de forma que, quando o estudante se deparar com determinado problema, possa inferir qual padrão ou combinação de padrões de programação pode ser utilizado na resolução do mesmo.

As categorias de padrões mais citadas em artigos da área estão descritas em BERGIN [1999] com relação a padrões de seleção e em ASTRACHAN & WALLINGFORD [1998], que descreve os padrões de repetição. BERGIN [1999] se preocupa também em apresentar, além dos padrões de seleção propriamente ditos, padrões de estratégia de resolução de problemas, padrões auxiliares e padrões estilísticos. Todavia, neste estudo, foram abordados apenas quatro padrões de seleção que serão descritos posteriormente.

3. Desenvolvimento do trabalho

A presente Seção demonstra os detalhes de como o trabalho foi desenvolvido. Na Seção 3.1, demonstramos o ambiente escolar, destacando como é ofertado o Ensino Médio Integrado no IFMT, o perfil dos estudantes que participaram do estudo, bem como alguns problemas encontrados no trabalho. A Seção 3.2 pontua os quatro padrões de programação abordados no trabalho. A Seção 3.3 discorre sobre o método de ensino utilizado discutindo alguns resultados obtidos.

3.1 - Caracterizações do ambiente escolar

Na modalidade Ensino Médio Integrado, os alunos do IFMT frequentam as aulas das disciplinas do núcleo comum, tais como Língua Portuguesa, Matemática, História, Geografia dentre outras no período da manhã, das 07:00 às 12:30, sendo que as disciplinas do núcleo profissionalizante são cursadas no período da tarde. Nesta modalidade, o *campus* Barra do Garças oferece os cursos de Técnico em Alimentos, Técnico em Comércio, Técnico em Controle Ambiental e Técnico em Informática.

O trabalho foi desenvolvido com alunos do primeiro ano do Ensino Médio, nos meses de abril, maior e parte de junho, o que corresponde ao primeiro bimestre de 2012. A classe, inicialmente, era composta por 39 estudantes, na faixa etária entre 13 a 15 anos, sendo que a distribuição por sexo se dá em 23 alunos do sexo masculino e 17 alunas do sexo feminino. A maioria, num total de 25 estudantes, é egressa de escolas públicas e o restante de escolas privadas. As aulas da disciplina Algoritmos e Lógica de Programação aconteceram às terças-feiras, das 14:30 às 17:30, com intervalo de 20 minutos iniciando às 16:00. A disciplina é ofertada anualmente para o primeiro ano do Ensino Médio, com 120 horas/aula de duração, sendo 30 horas/aula por bimestre letivo.

3.2 - Padrões de Seleção Utilizados no Trabalho

Para a apresentação e discussão dos padrões de programação com os alunos foi desenvolvido material instrucional baseado nos padrões de programação apresentados por BERGIN [1999]. Os padrões de programação apresentados e aplicados neste trabalho, bem como o foco do material instrucional, são apresentados a seguir.

3.2.1 - Seleção única

Este padrão de seleção é utilizado quando existe a necessidade de testar se uma condição é apropriada dependendo da condição de teste. No exemplo abaixo, o comando $c = a + b$ só será executado se a condição de teste for verdadeira, caso contrário o programa continuará sua execução.

```
if (a > b) {
    c = a + b;
}
```

3.2.2 - Escolha alternativa

Este padrão é utilizado quando se tem exatamente duas condições possíveis no teste lógico relacional. Ao contrário do padrão seleção única este padrão contém uma segunda parte, que é executada caso o primeiro teste for falso. No exemplo abaixo, caso a variável a seja menor ou igual a b , será executada a parte do *else*, que é $c = a + b$.

```
if (a > b) {
    c = b - a;
} else {
    c = a + b;
}
```

3.2.3 - Escolha sequencial

O padrão escolha sequencial é utilizado quando existem mais de duas condições possíveis no teste, desde que haja a necessidade de se fazer testes relacionais entre uma ou mais variáveis e um valor, ou entre uma ou mais variáveis com outras variáveis. O exemplo abaixo contém o trecho de um programa que verifica que tipo de triângulo é formado, dada as entradas realizadas pelo usuário. O uso do padrão escolha sequencial é indicado, pois existem mais de dois testes a serem realizados e existem testes entre variáveis.

```
if (a == 0 || b == 0 || c == 0) {
    cout << "Nao eh um triangulo ";
} else if (a > b+c || b > a+c || c > a+b) {
    cout << "Nao eh um triangulo ";
} else if (a == b && b == c) {
    cout << "Triangulo equilatero ";
} else if (a == b || b == c || a == c) {
    cout << "Triangulo isosceles ";
} else {
    cout << "Triangulo escaleno ";
}
```

3.2.3 - Série de Possibilidades

Este padrão é utilizado quando existem várias condições possíveis no teste, desde que não seja necessário algum teste relacional entre a variável de teste com algum valor ou da uma variável de teste com outra variável. O exemplo a seguir utiliza o padrão Série de Possibilidades, para fazer um programa que simula uma calculadora contendo as quatro operações básicas. Cada possibilidade possui uma instrução *break*, que interrompe a execução do comando relacionado ao padrão.

```

switch (opc) {
    case 1:
        result = a + b;
        break;
    case 2:
        result = a - b;
        break;
    case 3:
        result = a * b;
        break;
    case 4:
        result = a / b;
        break;
    default:
        cout << "Escolha invalida";
}

```

3.3 - Método de Ensino Utilizado

Nesta Seção, descrevemos o a aplicação do método de ensino utilizado no trabalho que resultou em um experimento bem sucedido. O experimento foi dividido em duas partes: apresentação e discussão dos padrões de programação e aplicação de uma lista de exercícios a serem codificados em uma linguagem de programação, explorando os padrões apresentados.

3.3.1 - Apresentação e discussão dos padrões de programação

Primeiramente, os alunos foram apresentados aos padrões, tentando manter sempre um paralelo entre o que eles vivem diariamente. Por exemplo, o padrão da escola é não utilizar boné ou bermuda em sala de aula. Ou o padrão das carteiras do laboratório é a cor azul. Assim, é solicitado que eles apontem o que está fora dos padrões dentro de sala de aula. As apresentações dos padrões desta forma têm como objetivo deixar a aula mais leve, buscando que os exemplos sejam retirados do mundo no qual o aluno vive, não o mundo abstrato da programação de computadores.

Os alunos trabalharam em duplas nos computadores do laboratório. Conforme McDOWELL [2002], estudantes que estudam programação em pares produzem melhores programas que alunos que estudam individualmente. Parte do trabalho também foi feito em ambientes fora do laboratório de informática.

Após a discussão dos padrões, os mesmos são mostrados em forma de linguagem de programação. A linguagem de programação utilizada nas aulas foi a linguagem C++ e os exemplos foram mostrados utilizando a IDE Dev C++ versão 4.9.2.

3.3.2 - Implementação da Lista de Exercícios

Após a apresentação dos padrões, foi lançada uma lista contendo 10 exercícios, com diferentes graus de dificuldade e foi disponibilizada no grupo virtual *IFMT-Informática* na rede social *Facebook*. Para cada exercício da lista, foi solicitado que, através de um comentário no cabeçalho do código de implementação, o estudante responda qual padrão está sendo utilizado na resolução do problema bem como o porquê da utilização do referido padrão. Assim, para cada exercício, o estudante poderá refletir, tanto individualmente, quanto em dupla, como se chegou à resolução do problema. Dessa forma, ele pode utilizar a resolução de um problema através de um padrão, para resolver outro problema semelhante, utilizando a mesma estrutura.

Com o objetivo de diagnosticar qual padrão os alunos tiveram mais dificuldades de compreender para a resolução da lista de exercícios, foi lançada uma enquete via internet, onde os estudantes tinham que responder, dentre os padrões utilizados em sala, baseando-se na resolução da lista de exercícios, qual foi o que o ele teve mais

difficultades para utilizar. A enquete foi criada no *Wordpress*, com bloqueios através de *cookies* e endereço IP, para garantir que um aluno participasse mais de uma vez da enquete. Dos 39 alunos, 15 responderam a enquete on-line. O padrão considerado de maior dificuldade de compreensão foi o Série de Possibilidades, com 71,43% (11 votos). A tabela a seguir, mostra o resultado da enquete.

Tabela 1: resultado da enquete on-line

Padrão	Votos	Percentual
Seleção única	0	0.00%
Escolha alternativa	1	7.14%
Escolha sequencial	3	21.43%
Série de possibilidades	11	71.43%

Considerando minha experiência prévia como docente, ao basear as soluções dos problemas em padrões de programação, os estudantes conseguiram soluções mais rápidas para os problemas dados. Muitos alunos, apesar da solução do problema estar correta, confundiram alguns comandos, especialmente os de relação entre uma variável e outra, bem como uma variável e um valor (igual, atribuição, diferente, maior, maior ou igual, menor, menor ou igual).

4. Conclusão e Trabalhos Futuros

A abordagem utilizando-se padrões de programação contextualizados no mundo real mostrou-se promissora, pois desta forma, o estudante consegue, através de modelos previamente estudados, resolver problemas de programação com maior eficiência com relação ao método tradicional de ensino. Interpretar o problema, buscando sua resolução através de soluções já conhecidas e testadas (padrões), parece ser uma abordagem eficaz no ensino de algoritmos para estudantes iniciantes em programação no Ensino Médio.

Trabalhos futuros serão elaborados utilizando padrões e jogos com materiais concretos em ambientes não formais para o ensino de programação de computadores. Outros trabalhos futuros poderão envolver estudos de combinação de padrões de seleção e padrões de programação de repetição.

5. Agradecimentos

Agradecemos o auxílio financeiro da Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG), e a cessão do espaço para a realização da pesquisa pelo Instituto Federal de Mato Grosso (IFMT), *campus* Barra do Garças.

6. Referências

ASTRACHAN, O. , WALLINGFORD E. (1998). “Loop Patterns”, <http://www.cs.duke.edu/~ola/patterns/plopd/loops.html>, Junho 2012.

- BARROS, L. N., MOTA, A. P. S., DELGADO, V. K., MATSUMOTO, P. M. (2005) “*A Tool for Programming Learning with Pedagogical Patterns*”, In: Proceedings of the 2005 OOPSLA - Workshop on Eclipse Technology eXchange, ACM Press, p. 125-129.
- BORGES, M. A. F. (2000) “*Avaliação de uma Metodologia Alternativa para a Aprendizagem de Programação*”. In: VIII Workshop de Educação em Computação – WEI 2000. Curitiba, PR, Brasil.
- BERGIN, J. (1999). “*Patterns for Selection Version 4*”, <http://csis.pace.edu/~bergin/patterns/PatternsV4.html>, Junho 2012.
- CLANCY, J. e LINN, M. C. (1999) “*Patterns and Pedagogy*”, In: Proceedings of the 30th SIGCSE Technical Symposium on Computer Science Education, ACM Press, p. 37-42.
- HABERMAN, B. e MULLER, O. (2008) “*Teaching Abstraction to Novices: Pattern-Based and ADT-Based Problem-Solving Processes*”, In: 38th ASEE/IEEE Frontiers in Education Conference, IEEE, p. F1C-7-F1C12.
- IFMT (2011). “*Exame de Seleção 2012/1 - Ensino Médio Integrado - Edital N° 61/2011*”, http://selecao.ifmt.edu.br/download.aspx?cod_arquivo_download=3864, Junho, 2012.
- JOHNSON, W. L. e SOLOWAY, E. (1985) “*PROUST: Knowledge-based program understanding*”. In: IEEE, Transactions on Software Engineering, IEEE, p. 369-380.
- MINISTÉRIO DA EDUCAÇÃO. “*Institutos Federais – Uma Conquista de Todos os Brasileiros*”, http://portal.mec.gov.br/dmdocuments/folheto_setec.pdf, Junho 2012.
- MULLER, O. (2005) “*Pattern Oriented Instruction and the Enhancement of Analogical Reasoning*”, In: Proceedings of the first international workshop on Computing education research (ICER), ACM Press, p. 57-67.
- MULLER, O., HABERMAN, B.,GINAT, D. (2007) “*Pattern-Oriented Instruction and its Influence on Problem Decomposition and Solution Construction*”, In: Proceedings of the 12th Annual SIGCSE - Conference on Innovation and Technology in Computer Science Education (ITiCSE'07), ACM Press, p. 151-155.
- PEREIRA JÚNIOR, J. C. R., RAPKIEWICZ, C.E. (2004) “*O Processo de Ensino-Aprendizagem de Programação: Uma Visão Crítica da Pesquisa no Brasil*”, In: WEI - Workshop de Informática na Educação 2004.
- PEREIRA JÚNIOR, J. C. R., RAPKIEWICZ, C. E., DELGADO, C., XEXEO, J. A. M. (2005). “*Ensino de Algoritmos e Programação: Uma Experiência no Nível Médio*”. In: XIII Workshop de Educação em Computação - SBC2005, 2005, São Leopoldo.
- PORTER, R. e CALDER, P. “*Pattens in Learning to Program – An Experiment?*”(2004), In: Proceedings of the Sixth Australasian Conference on Computing Education, ACM Press, p. 241-246.
- SOLOWAY, E. (1986) “*Learning to Program = Learning to Construct Mechanisms and Explanations*”, In: Communications of the ACM, ACM Press, p. 850-858.
- WINSLOW, L. E. (1996) “*Programming Pedagogy – A Psychological Overview*”, In: ACM SIGCSE Bulletin, ACM Press, p. 17-22.

Modelagem de Vídeo MPEG-4 Utilizando Cascata Multifractal com Multiplicadores Autorregressivos

Flávio Geraldo Coelho Rocha, Flávio Henrique Teles Vieira e Douglas Borges de Oliveira Filho

Escola de Engenharia Elétrica e de Computação (EEEC/UFG)
Av. Universitária, nº.1488, Qd. 86, Bloco A
74605-010, Goiânia, GO, Brasil

flavio.geraldo@gmail.com, flavio@eee.ufg.br, douglasfilho@gmail.com

Abstract. *In this paper we propose a multifractal model for MPEG-4 video traffic based on a multiplicative cascade. In this proposed model we model the multipliers distribution for each scale by autoregressive processes. We also verify the model performance in capturing the real traffic traces characteristics compared to other multifractal models.*

Resumo. *Neste artigo, propomos um modelo multifractal para tráfego de vídeo MPEG-4 baseado em uma cascata multiplicativa. Nesta proposta modelamos a distribuição dos multiplicadores através de processos autoregressivos a cada escala. Verificamos também o desempenho do modelo em capturar as características das séries reais de tráfego em comparação a outros modelos multifractais.*

1. Introdução

Aplicações que envolvem tráfego de vídeo são amplamente utilizadas em redes de telecomunicações. O crescente número de aplicações multimídia com requisitos de qualidade de serviço (QoS) sugere que grande parte do tráfego da rede mundial de computadores será composta por tráfego de vídeo [Fei e Zhimei 2003]. Mesmo utilizando-se de técnicas de compressão como o MPEG-4, o tráfego de vídeo demanda uma grande largura de banda, principalmente em aplicações de tempo real, cada vez mais requeridas pelos usuários [Wang e Qiu 2005].

A modelagem de vídeo não é um tópico recente. Para um melhor dimensionamento, utilização dos recursos e compreensão dos dados que trafegarão na rede, a modelagem de dados de vídeo tem sido um tópico de pesquisas constante. Inicialmente, modelos baseados em cadeias de Markov, assim como processos autoregressivos foram desenvolvidos [Heyman et al. 1992],[Maglaris et al.1988]. No entanto, estes modelos eram capazes apenas de capturar as características de curta dependência entre as amostras. Em outras palavras, um processo com curta dependência significa que sua função de autocorrelação que mede essa dependência, apresenta um rápido decaimento (exponencial) à medida que se considera amostras mais distantes no tempo. Quando características de longa dependência foram consideradas, modelos, como por exemplo, FGN [Huang et al. 1995] e FARIMA [Cruz et al. 1998] foram aplicados. Pesquisas recentes têm mostrado que o tráfego de vídeo possui propriedades

que vão além daquelas relativas a simples processos auto-similares. Longa-dependência entre as amostras e tráfego em rajadas persistentes nas mais diversas escalas têm sido observadas [Feldman et al. 1998], [Wang e Qiu 2006]. Neste sentido, modelos multifractais proporcionam melhores resultados em termos de desempenho de modelagem de tráfego [Riedi et al. 1999],[Krishna et al. 2003],[Vieira e Ling 2008].

Uma vez modelado o tráfego de rede, um projetista poderá valer-se do entendimento do comportamento do tráfego para, posteriormente, tomar decisões com base nas características encontradas, por exemplo dimensionar a banda realmente necessária para a transmissão dos dados. Além disso, modelos de tráfego podem ser utilizados em algoritmos para alocação eficiente de recursos e controle de taxa de transmissão em tempo real [Ouyang et al. 2006],[Aquino e Barria 2006].

Este trabalho tem o objetivo de propor um modelo baseado em cascata multifractal para modelagem do tráfego de vídeo de taxa variável (VBR- *Variable Bit Rate*) que utiliza MPEG-4 como técnica de compressão. O modelo proposto foi chamado de CMAM (Cascata com Modelagem Autorregressiva para os Multiplicadores). Testes de desempenho foram realizados e confirmaram o bom desempenho do modelo em comparação a outros modelos multifractais. O foco deste artigo é o de avaliar o modelo CMAM em termos de desempenho de modelagem, inclusive para séries de vídeo diferentes das apresentadas em [Rocha e Vieira 2011]. Além disso, o objetivo principal do artigo [Rocha e Vieira 2011] é o de utilizar o modelo como entrada para um esquema de controle de taxa de tráfego de vídeo que atualiza parâmetros de um codificador de vídeo de forma adaptativa.

O artigo está organizado da seguinte forma: na seção 2, discursamos sobre princípios de modelos multifractais baseados em cascata multiplicativa. Na seção 3, apresentamos o modelo multifractal CMAM proposto. Na seção 4, revemos alguns conceitos e características da modelagem de vídeo MPEG-4 e são mostrados os testes realizados para validar o modelo multifractal CMAM. Por fim, na seção 5, concluímos.

2. Análise Multifractal

A descrição do comportamento local de medidas e funções em uma forma geométrica e estatística é de interesse da análise multifractal [Park e Willinger 2000]. Na análise multifractal verifica-se o comportamento em escala de momentos estatísticos dos processos para estimar suas regularidades locais [Feldman et al. 1998],[Riedi et al. 1999]. Através de ferramentas da análise multifractal algumas propriedades encontradas em processos reais podem ser verificadas. O tráfego de redes, por exemplo, ao ser considerado multifractal significa que possui uma estrutura de forte dependência inerente entre as amostras, com incidência de rajadas em várias escalas [Riedi et al. 1999]. Estas características podem degradar o desempenho de rede em relação a fluxos de tráfego Gaussianos e de curta-dependência [Riedi et al. 1999],[Vieira e Ling 2008],[Park e Willinger 2000].

2.1. Cascatas Multiplicativas

O conceito de cascata multiplicativa nos remete ao conceito de cascata binomial. A cascata binomial é um método de se obter um processo multifractal que consiste de um processo iterativo no intervalo compacto $[0,1]$. Ao se permitir que os multiplicadores da cascata sejam variáveis aleatórias independentes em $[0,1]$ com densidade de

probabilidade $f_R(X)$, obtém-se uma estrutura geral chamada cascata multiplicativa em que os multiplicadores são valores fixos.

2.2. Estimação da Densidade de Probabilidade dos Multiplicadores

Seja X^N o processo de tráfego no estágio N da cascata. A série de tráfego no estágio $(N-1)$ da cascata pode ser obtida agregando valores consecutivos do estágio N em blocos não-sobrepostos de tamanho 2. De forma análoga, dada a série na escala $(N-j)$, X_i^{N-j} onde $(i=1, \dots, 2^{N-j})$, obtemos os dados na escala $(N-j-1)$ pela soma consecutiva dos valores do estágio $(N-j)$ da seguinte forma:

$$X_i^{N-j-1} = X_{2i-1}^{N-j} + X_{2i}^{N-j} \quad (1)$$

para $i=1, \dots, 2^{N-j-1}$. Este procedimento termina quando a agregação dos valores forma apenas um ponto na última escala da cascata. Uma estimativa $r_j^{(i)}$ dos multiplicadores da cascata pode ser obtida pela equação [Feldman et al. 1998]:

$$r_j^{(i)} = X_i^{N-j} / X_{2i-1}^{N-j-1} \quad (2)$$

para $i=1, \dots, 2^{N-j-1}$. Podemos considerar $r_j^{(i)}$ como sendo amostras da distribuição $f_{R_j}(r)$ dos multiplicadores no estágio j . A distribuição dos multiplicadores na escala j pode ser obtida pelos histogramas de $r_j^{(i)}$. O modelo multifractal VVGM [Krishna et al. 2003], por exemplo, é uma cascata multiplicativa que aproxima os histogramas obtidos por gaussianas. Ou seja, neste modelo assume-se que a distribuição dos multiplicadores $f_{R_j}(r)$ é gaussiana centrada em 0.5 e com variâncias que mudam a cada escala. Essas variâncias são estimadas a partir de histogramas [Krishna et al. 2003].

3. Cascata com Modelagem Autorregressiva para os Multiplicadores (CMAM)

3.1. Modelo CMAM

Nesta seção, apresentamos o algoritmo de síntese do processo CMAM. O algoritmo de síntese de tráfego de vídeo proposto faz uso do valor agregado de tráfego na última escala da cascata e da densidade de probabilidade dos multiplicadores estimada através de processos autoregressivos.

A partir das equações (1) e (2) pode-se representar a estrutura de construção da cascata multiplicativa, descrita na seção 2, através da Figura 1:

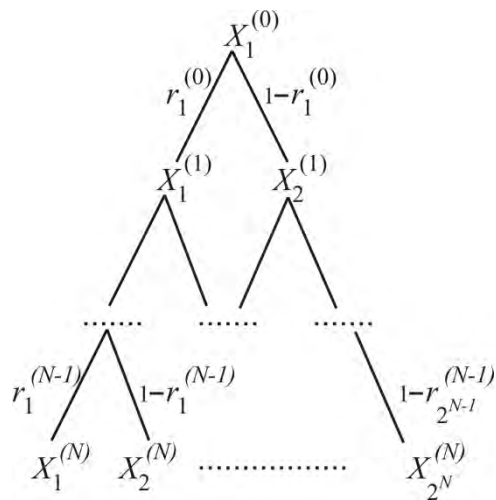


Figura 1. Processo de construção da cascata multiplicativa

Assim, o algoritmo de síntese do processo CMAM consiste basicamente dos seguintes passos:

Algoritmo 1

- 1) Calcula-se o valor agregado da série de tráfego obtido na última escala N ;
- 2) Estimam-se os multiplicadores usando as equações (1) e (2);
- 3) Calcula-se, através do método dos mínimos quadrados apresentado em [Neumaier e Schneider 2001], os parâmetros do modelo de regressão (processo autorregressivo) para cada escala da cascata;
- 4) Estima-se a distribuição dos multiplicadores em cada escala através de um processo autorregressivo;
- 5) Com os multiplicadores disponíveis, gera-se amostras de um processo multifractal por meio do procedimento descrito na seção 2.

4. Modelagem de Vídeo

Por existirem semelhanças entre os quadros sucessivos gerados a partir de uma mesma cena, o MPEG-4 é uma técnica de compressão que utiliza codificação tanto de redundância espacial (*intra-frames*) quanto codificação de redundância temporal (*inter-frames*) [ISO/IEC 2000]. No MPEG-4 existem três tipos diferentes de quadros, os quadros do tipo I, os quadros do tipo P e os quadros do tipo B. Os quadros do tipo I são os quadros mais importantes, maiores em bits e que carregam o maior número de informações importantes, estes quadros não dependem de nenhum outro. Já os quadros do tipo P dependem de quadros tipo I ou tipo P anteriores, enquanto os quadros do tipo B dependem de quadros anteriores e sucessivos dos tipos I e P.

Os quadros dos tipos I, P e B são organizados em grupos chamados de GoPs (*Group of Pictures* – Grupo de imagens), cada GoP é composto por um quadro do tipo I, de um certo número de quadros do tipo P e opcionalmente de alguns quadros do tipo B localizados entre os quadros I e P, e entre quadros do tipo P consecutivos. Uma série de GoPs compõe uma série de vídeo. Desta forma os quadros do tipo I são chamados de *intra-frames* e os quadros dos tipos P e B são chamados de *inter-frames*.

Com a finalidade de gerar séries sintéticas de tráfego de vídeo MPEG-4 foi utilizado o modelo multifractal baseado em cascata multiplicativa CMAM, apresentado na seção anterior, e sintetizado a partir do algoritmo 1. Um modelo de tráfego multifractal foi utilizado por já ter sido mostrado que o tráfego de vídeo MPEG-4 possui características multifractais, tais como diferentes leis de escala, decaimento lento da função de autocorrelação e rajadas que persistem para diferentes escalas [Fei e Zhimei 2003],[Wang e Qiu 2005].

4.1. Comparação de Estatísticas do Modelo Multifractal Proposto em Relação a outros Modelos Multifractais

Com o objetivo de comparar as características estatísticas do modelo CMAM e do tráfego de vídeo real, assim como de outros modelos multifractais de tráfego, são apresentados resultados e testes estatísticos como: média, variância e função densidade de probabilidade (*pdf-probability density function*).

Neste artigo, comparamos o desempenho de nosso modelo com o do modelo multifractal VVGM (*Variable Variance Gaussian Model*) [Krishna et al. 2003], que se destaca por sua simplicidade, e também o comparamos com o MWM (*Multifractal Wavelet Model*) [Riedi et al. 1999], que apresenta um ótimo desempenho para o tráfego multifractal [Riedi et al. 1999],[Vieira e Ling 2008],[Ribeiro et al. 2000].

Como exemplos, apresentamos nesta seção os resultados obtidos para duas séries de vídeo, sendo elas: “*Silence of the Lambs*” (O Silêncio dos Inocentes) e “*Lord of the Rings 3*” (Senhor dos Anéis 3), chamados neste trabalho de S.LAMBS e LORD3, respectivamente. Estes vídeos utilizam compressão MPEG-4 parte 2, resolução de 352 x 288 pixels e para análise apresentada foram considerados os 32768 primeiros quadros das séries. A série de vídeo S.LAMBS possui taxa de 30 quadros por segundo e agrupamento de 16 quadros por GoP, ordenados da seguinte forma (IBBBPBBBPBBBPBBB – G16B3). A série de vídeo LORD3 possui taxa de 25 quadros por segundo e agrupamento de 12 quadros por GoP ordenados da seguinte forma (IBBPBBBPBBBPBB – G12B2). Mais informações sobre as características das séries de vídeo podem ser encontradas em [Seeling et al. 2004],[Rocha e Vieira 2011].

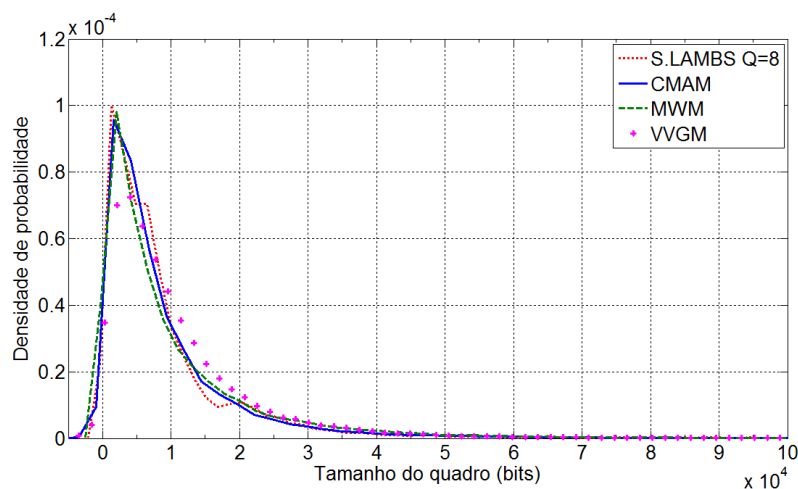


Figura 2 – Função densidade de probabilidade (*pdf*) – S.LAMBS

A Figura 2 apresenta as funções de densidade de probabilidade reais e sintéticas da série de vídeo S.LAMBS. Observa-se que esta série possui uma tendência a ser

descrita por uma função lognormal e que o modelo CMAM foi capaz de representar bem este comportamento. Em relação aos modelos VVGM e MWM, o modelo CMAM apresentou desempenho comparável aos desempenhos dos demais modelos e em alguns casos, como na Figura 2, apresentou desempenho superior ao VVGM. Resultados semelhantes foram obtidos para outras séries de vídeo. A análise estatística das séries de vídeo e as implementações dos códigos para a modelagem (VVGM, MW, CMAM) das séries foram realizadas no *software* Matlab.

Tabela 1. Média e variância – série de quadros – S.LAMBS

Quadros (bits)	Média	Erro	Variância	Erro
S.LAMBS	$9,411.10^3$	-	$1,659.10^8$	-
VVGM	$1,006.10^4$	6,84%	$1,033.10^8$	37,77%
MWM	$1,126.10^4$	19,64%	$1,856.10^8$	11,83%
CMAM	$9,280.10^3$	1,40%	$1,456.10^8$	12,26%

A Tabela 1 apresenta estatísticas do vídeo S.LAMBS. Por meio da Tabela 1 observa-se que a média obtida com o modelo CMAM proposto foi a que mais se aproximou da série real e que a variância encontrada com o CMAM é comparável aquela obtida com o MWM e mais próxima do real que o modelo VVGM. Isto pode ser observado através da coluna erro, que apresenta o erro relativo das estatísticas das séries sintéticas geradas em relação às séries reais.

A Tabela 2 apresenta as estatísticas do vídeo LORD3. Nesta tabela é possível observar que a média obtida para o modelo CMAM está muito próxima da série real e que a variância encontrada está mais próxima do real que o importante modelo MWM [Riedi et al. 1999].

Tabela 2. Média e variância – série de quadros – LORD3

Quadros (bits)	Média	Erro	Variância	Erro
LORD3	$2,872.10^3$	-	$9,362.10^6$	-
VVGM	$3,087.10^3$	7,52%	$1,082.10^7$	15,60%
MWM	$2,416.10^3$	16,03%	$6,296.10^6$	32,75%
CMAM	$2,844.10^3$	0,96%	$1,172.10^7$	25,21%

5. Conclusão

Pode-se concluir que o modelo multifractal proposto consegue capturar com eficiência as características do tráfego de vídeo VBR MPEG-4. Essa afirmação é confirmada pelos testes realizados.

Neste trabalho, propomos uma modelagem da distribuição dos multiplicadores através de processos autoregressivos para cada estágio da cascata, logo a média dos multiplicadores varia para cada escala; diferente do que ocorre no VVGM [Krishna et al. 2003]. Esta característica do modelo proposto permite que um desempenho eficiente de modelagem seja obtido. Os testes de desempenho de modelagem revelam que os

resultados do modelo CMAM são comparáveis e, na maioria dos casos, superiores ao do modelo MWM [Riedi et al. 1999].

Referências

- Aquino, V. A. e Barria, J. A. (2006) "Multiresolution FIR neural-network-based learning algorithm applied to network traffic prediction". IEEE Transactions on Systems, Man and Cybernetics-C, vol. 36, no.2, pp.208-220.
- Cruz, L. J., Pallares, E. e Alins, J. J. (1998) "Self-similar traffic generation using a fractional ARIMA model. Application to the VBR MPEG video traffic," in Telecommunications Symposium, 1998. ITS'98 Proceedings. SBT/IEEE International, vol. 1, pp. 102-107.
- Fei, H. e Zhimei, W. (2003) "Multifractal analysis and model of the MPEG-4 video traffic," in Performance, Computing, and Communications Conf., vol. 9-11, pp. 463-467.
- Feldmann, A., Gilbert, A. C. e Willinger, W. (1998) "Data networks as cascades: Investigating the multifractal nature of Internet WAN traffic". ACM/SIGCOMM'98, Vancouver. pp. 25-38.
- Heyman, D., Tabatabai, A. e Lakshman, T. V. (1992) "Statistical analysis and simulation study of video conference traffic in ATM networks," IEEE Trans. Circuits and Systems for Video Technology, vol. 2, no. 1, pp.49-59.
- Huang, C., Devetsikiotis, M., Lambadaris, I. e Kaye, A. R. (1995) "Fast Simulation for Self-Similar Traffic in ATM Networks." In Proc. IEEE ICC, pp. 438-444.
- ISO/IEC/JTC1/SC29/WG11. (2000) "Overview of the MPEG-4 standard".
- ISO/IEC/JTC1/SC29/WG11 MPEG Test Model 2 (1992) "Coded Representation of Picture and Audio Information".
- Krishna, M. P., Gadre, V. M. e Dessay, U. B. (2003) "Multifractal based network traffic modeling". Kluwer Academic Publishers.
- Liu, H. H., Hsu, P. L. (2006) "Design and simulation of adaptive fuzzy control on the traffic network". In International Joint Conference SICE-ICASE, pp. 4961-4966.
- Maglaris, B., Anastassiou, D., Sen, P., Karlsson, G. e Robbins, J. (1988) "Performance Models of Statistical Multiplexing in Packet Video Communications", vol. 36, no. 7, pp. 834-844.
- Neumaier, A. e Schneider, T. (2001) "Estimation of parameters and eigenmodes of multivariate autoregressive models". ACM Trans. Math. Softw. 27, 1 (Mar.), 27-57.
- Ouyang, Y. C., Yang, C. W. e Lian, W. S. (2006). "Neural networks based variable bit rate traffic prediction for traffic control using multiple leaky bucket". Journal of High Speed Networks. vol. 15, no.2, pp.11-122.
- Park, K. e Willinger, W. (2000) "Self-similar Network Traffic and Performance Evaluation". John Wiley and Sons, New York.
- Ribeiro, V.J., Riedi, R.H., Crouse, M.S. e Baraniuk, R.G.(2000) Multiscale queueing analysis of long-range dependent traffic, Proc. IEEE Infocom.

- Riedi, R. H., Crouse, M. S., Ribeiro, V. J. e Baraniuk, R. G.(1999) “A multifractal wavelet model with application to network traffic”. IEEE Trans. on Information Theory, vol. 45, no.3, pp. 992–1018.
- Rocha, F.G.C. e Vieira, F.H.T. (2011) “Modelagem e Controle da Taxa de Vídeo MPEG Utilizando Cascata Multifractal com Multiplicadores Autorregressivos, XXIX Simpósio Brasileiro de Telecomunicações, SBrT, Curitiba PR.
- Seeling, P., Reisslein, M., Kulapala, B. (2004) “Network Performance Evaluation Using Frame Size and Quality Traces of Single-Layer and Two-Layer Video: A Tutorial”. IEEE Communications Surveys and Tutorials, Vol. 6, No. 2, Pages 58-78.
- Vieira, F. H. T. e Ling, L. L. (2008) “Modelagem de Tráfego de Redes Utilizando Cascata Multifractal Generalizada”. RITA, Vol.15, No. 2.
- Wang, S. H. e Qiu, Z. D. (2005) “A Novel Multifractal Model of MPEG-4 Video Traffic”. Proceedings of ISCIT2005.
- Wang, S. H. e Qiu, Z. D., (2006) “Multifractal Analysis and Prediction of VBR Video Traffic”. 6th International Conference on ITS Telecommunications Proceedings.
- Zainaldin, A., Lambadaris, I., e Nandy, B. (2008) “Adaptive rate control low bit-rate video transmission over wireless Zigbee networks”. Proceedings of ICC2008.

Caracterização de QoS para Conteúdo Multimídia em Rede VPN utilizando Rede Neural MLP

Paula Letícia S. Lima, Ananias Pereira Neto

Centro de Ciências Exatas e Tecnologia (CCET) – Curso de Bacharelado em Ciência da Computação – Universidade da Amazônia (UNAMA)

{palelima}@gmail.com, apneto@unama.br

Abstract. *This paper shows a classification method of QoS in a VPN network in dissemination of media content. This method uses the MLP (Multilayer Perceptron) neural network to establish relation between the QoS metrics, such as Jitter, Throughput, Packets Lost and Data Transferred. The neural network classifies QoS ambience in High, Medium and Low in dissemination of content.*

Resumo. *Este artigo apresenta um método de classificação de QoS em rede VPN (Virtual Private Network) na disseminação de conteúdo multimídia. Esse método utiliza a rede neural MLP (Multilayer Perceptron) para estabelecer relação entre as métricas de QoS, tais como Jitter, Throughput, Perda de Pacotes e Dados Transferidos. A rede neural classifica o ambiente de QoS em Alto, Médio e Baixo na disseminação de conteúdo.*

1. Introdução

Para caracterizar um ambiente de QoS (*Quality of Service*), de um determinado conteúdo multimídia, são exigidas métricas convencionais entre valores máximos e mínimos. O termo QoS refere-se à capacidade da rede garantir um certo nível de desempenho para o tráfego multimídia. QoS é um conceito técnico medido, expresso e compreendido no nível de rede e geralmente tem pouco significado para o usuário.

Segundo [GOMES, et. al. 2009] diferentes soluções foram desenvolvidas para proporcionar QoS. Para a percepção humana é complexo classificar os diversos parâmetros de QoS, fazendo necessário utilizar regras computacionais como as redes neurais artificiais. A Rede Neural Perceptron Multi-Camadas tem característica de tratar os parâmetros de QoS, pois utiliza de camadas intermediárias conhecidas como, camadas ocultas, trabalhando como reconhecedor desses parâmetros.

Inúmeros trabalhos procuram avaliar o ambiente de QoS em conteúdo multimídia e foram propostos na literatura, tais como, [Y. BAI 2011], [GÓMEZ 2010], [GUEDES 2009], poucos com a técnica de MLP (Multilayer Perceptron) para redes com fio. A proposta deste artigo é classificar o conteúdo multimídia em uma rede VPN utilizando a rede neural MLP, baseado em métricas convencionais, tais como, *Jitter*, *Throughput*, Perda de Pacotes e Quantidade de dados transferidos.

Este artigo está organizado em cinco seções. A seção 2 apresenta os parâmetros de QoS para classificação do nível da rede. Na seção 3, são apresentados conceitos relativos à Rede Neural MLP, bem como sua aplicabilidade na classificação de QoS em conteúdo multimídia. Na seção 4 apresenta o cenário em que foram desenvolvidos os

testes e obtidos os resultados a partir das métricas aplicadas ao experimento. A seção 5 apresenta as conclusões desenvolvidas nesse trabalho.

2. Classificação de QoS

As redes de computadores dispõem de uma gama de dados e aplicações em tempo real, tornando necessária a medição da qualidade dos serviços. Os parâmetros de QoS buscam tratar de prioridades, alocação de banda e controle de latência da rede. Porém, a maioria das redes não é segura e dotada de qualidade de serviço.

Quando se faz alguma transmissão na rede o pacote pode ou não chegar ao seu destinatário, sendo assim os parâmetros de QoS começam a ser visto como uma parte importante na implementação de uma rede. Segundo [GUEDES 2009], existem diversos parâmetros de QoS para caracterizar uma rede de computadores e a disseminação de conteúdo multimídia, tais parâmetros como *Throughput*, *Jitter*, Perda de Pacotes e Dados Transferidos.

Os parâmetros aplicados neste trabalho para a classificação do conteúdo multimídia em rede VPN foram determinados pela utilização de técnica de mineração de dados, usando a ferramenta Bayesware Discoverer [BAYESWARE, 2000]. Este software foi utilizado para determinar o valor mínimo e máximo de cada classificação dos parâmetros em Alto QoS, Médio QoS e Baixo QoS, utilizando métricas: Perda de Pacotes, *Jitter*, *Throughput* e Dados Transferidos.

3. Multilayer Perceptron

Uma RNA (Rede Neural Artificial) pode ser definida como sendo estrutura de rede, passível de implementação em dispositivos eletrônicos, composta por um número de unidades interconectadas (neurônios artificiais), os neurônios artificiais recebem essa denominação porque foram originados a partir de um modelo matemático de um neurônio biológico; a principal propriedade desta arquitetura é a sua capacidade de aproximação universal e de aprendizado [FERNANDES 2008].

A Multilayer Perceptron é uma extensão do Perceptron simples, capaz de trabalhar com problemas não-linearmente separáveis. Este avanço foi possível através da utilização de, pelo menos, uma camada entre a entrada e a saída. Nas redes MLP, cada camada tem sua função determinada. A camada intermediária enviar estímulos para camada de saída, essa camada realiza cálculos necessários para construir um padrão que será a resposta. A função das camadas intermediárias é extrair características. Seus pesos são um código de características apresentadas nos padrões de entrada permitindo que a rede crie sua representação, mais rica e complexa.

Neste trabalho, foram coletadas 150 amostras, usando quatro medidas na rede VPN, Dados Transferidos, Perda de Pacotes, *Throughput* e *Jitter*. Os dados foram divididos em dados de treino e de validação, além do desenvolvimento da rede MLP para estabelecer a relação entre as métricas variando o número de neurônios na camada escondida.

4. Experimentos e Resultados

Esta seção descreve as simulações realizadas neste artigo, assim como os resultados obtidos. Para isso, foram avaliadas métricas de desempenho em um cenário composto por quatro computadores Itautec com processadores Intel Celeron 1.73GHz e 1GB de memória RAM conectados a um servidor de mesma configuração. Tal cenário pode ser observado na Figura 1.

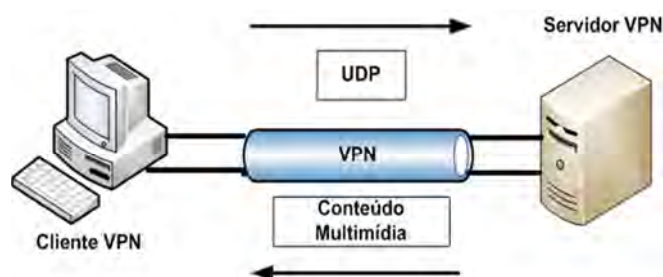


Figura 1. Cenário do experimento.

No experimento realizado foi implementada uma rede VPN e analisado o desempenho desta rede utilizando testes com tráfego multimídia e sem tráfego multimídia. As medições foram realizadas utilizando o software *iperf* [SCHRODER 2008] variando o número de clientes, com o objetivo de avaliar com maior precisão os resultados dos experimentos.

As estações clientes executavam um arquivo de vídeo que estava armazenado no servidor simultaneamente e a partir do *iperf* foram adicionados no tráfego da VPN datagramas UDP para análise do comportamento das métricas. As informações com o desempenho da rede chegavam ao servidor, no qual foram analisadas.

O Servidor VPN é responsável por realizar toda a distribuição das tarefas e o monitoramento do desempenho dos clientes. Neste sentido, ele realiza o monitoramento das falhas que podem ocorrer e o direcionamento da carga de processamento, caso haja alguma indisponibilidade, o que garante verificar o comportamento do VPN a partir das métricas de desempenho aplicadas.

Tabela 1. Medidas estatísticas para os parâmetros de QoS medidos no experimento.

	Média	Desvio Padrão	Variância	Intervalo de Confiança
Sem Tráfego Multimídia				
Dados Transferidos (MB)	0,706	0,401	0,161	0,706±0,064
Throughput (MB/s)	1,001	0,151	0,023	1,001±0,024
Jitter (ms)	4,410	1,183	1,400	4,410±0,189
Perda de Pacotes	387,473	284,893	81163,875	387,473±45,591
Com Tráfego Multimídia				
Dados Transferidos (MB)	0,726	0,451	0,203	0,726±0,072
Throughput (MB/s)	1,050	0,096	0,009	1,050±0,015
Jitter (ms)	4,542	1,413	1,997	4,542±0,226
Perda de Pacotes	401,567	320,811	102919,388	401,567±51,339

A Tabela 1 apresenta os dados da média, desvio padrão, variância e intervalo de confiança com e sem tráfego multimídia, considerando um nível de confiança de 95% para as métricas do experimento realizado.

A partir dos dados medidos no experimento, o banco de dados foi dividido em dados de treino e de validação, conforme as 50 medidas de amostra para os 3 tipos de classificações da rede que chamamos de Alto, Médio e Baixo QoS. Constituindo 35 padrões de classificação para treino e 15 padrões de classificação para validação. Após o tratamento dos padrões de treino e validação, desenvolveu-se a rede MLP para a classificação.

Na simulação da rede MLP foram realizados 50 treinos, variando o número de neurônios na camada escondida, com o intuito de determinar a melhor topologia baseado na análise da percentagem de acerto da rede MLP para os dados de treino e validação. A Tabela 2 apresenta os dados da rede MLP variando o número de neurônios na camada escondida, obtendo-se o erro para dados de treino e validação, além da porcentagem de erro para a rede MLP sem e com tráfego multimídia.

Tabela 2. Dados da Rede MLP sem e com tráfego multimídia.

	<i>Sem Tráfego Multimídia</i>			<i>Com Tráfego Multimídia</i>		
	RNA 1	RNA 2	RNA 3	RNA 1	RNA 2	RNA 3
Nº de Entradas	4	4	4	4	4	4
Nº de Neurônios na Camada Escondida	1	5	10	1	5	10
Erro para os Dados de Treino	0,2292	0,0429	0,0748	0,0770	0,0435	0,0291
Erro para os Dados de Validação	0,1742	0,1264	0,1486	0,0915	0,0473	0,0302
Porcentagem de Erro	20 %	13,33 %	17,78 %	15,56 %	4,44 %	2,22%

Nota-se no experimento, que o melhor caso sem tráfego multimídia foi a RNA 2, com porcentagem de erro de 13,33 %, com 5 neurônios na camada escondida e baixo índice de erro comparado ao pior caso, conforme os dados apresentados na Tabela 2. A rede MLP classificou 31,43 % em Alto QoS, 35,24 % em Médio QoS e 33,33 % em Baixo QoS nos dados de treino.

A Figura 2 (a) ilustra a distribuição da classificação de QoS sem tráfego multimídia para o melhor caso da rede MLP (RNA 2), neste gráfico observa-se que a classificação do QoS é maior em Médio QoS, comparando as métricas entre *Jitter* e Perda de Pacotes. Na Figura 2 (b) pode-se observar a relação da Perda de Pacotes e o *Throughput* baseada na porcentagem de erro do melhor caso (RNA 2) para a distribuição da classificação sem tráfego multimídia, a classificação após o treino apresentou uma distribuição com maior proporção em Médio QoS.

Para o melhor caso com tráfego multimídia, a rede MLP que apresentou menor porcentagem de erro foi a RNA 3, com 2,22 %, contendo 10 neurônios na camada escondida, conforme apresentado Tabela 2. Durante esse processo de treinamento é possível observar que a rede MLP (RNA 3) classificou 33,33 % em Alto QoS, 33,33 % em Médio QoS e 33,33 % em Baixo QoS nos dados de treino, uma distribuição uniforme entre os dados treinados entre Alto, Médio e Baixo QoS.

A distribuição da classificação de QoS com tráfego multimídia para o melhor caso da rede MLP (RNA 3), comparando as métricas entre *Jitter* e Perda de Pacotes é apresentada na Figura 3 (a), nota-se após o treinamento da rede MLP uma distribuição bem definida entre Alto, Médio e Baixo QoS, devido a rede apresentar um erro extremamente baixo comparado as demais redes treinadas. A Figura 3 (b) mostra a distribuição da classificação de QoS com tráfego multimídia para o melhor caso da rede MLP (RNA 3), comparando as métricas Perda de Pacotes e *Throughput*, observa-se novamente uma distribuição uniforme da classificação em Alto, Médio e Baixo QoS, como esperado pelos dados de treino e validação da rede MLP.

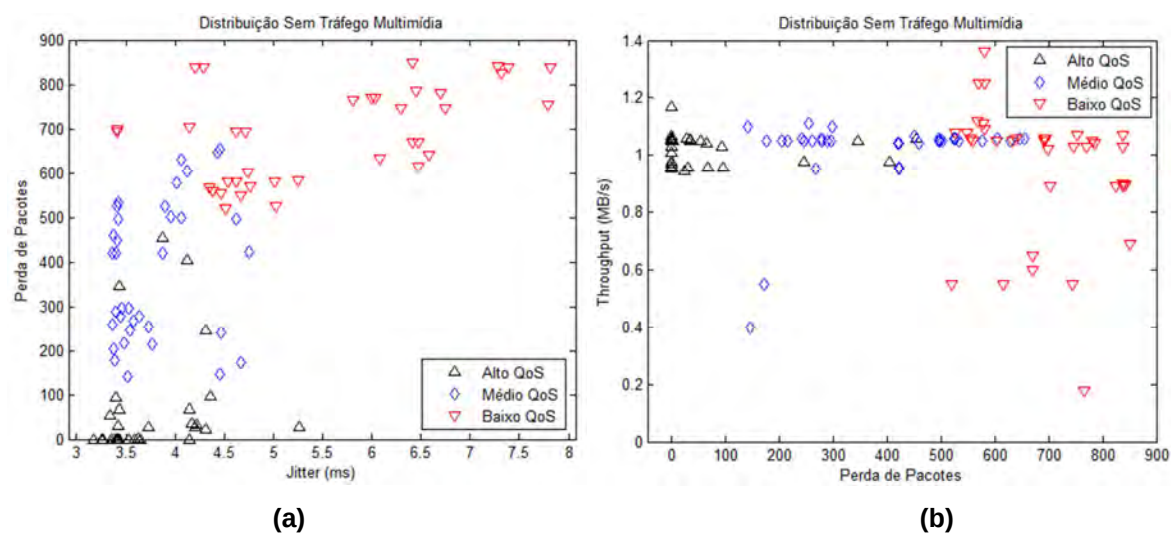


Figura 2. Rede MLP (RNA 2) sem tráfego multimídia: (a) Melhor caso Jitter x Perda de Pacotes (b) Melhor caso Perda de Pacotes x Throughput.

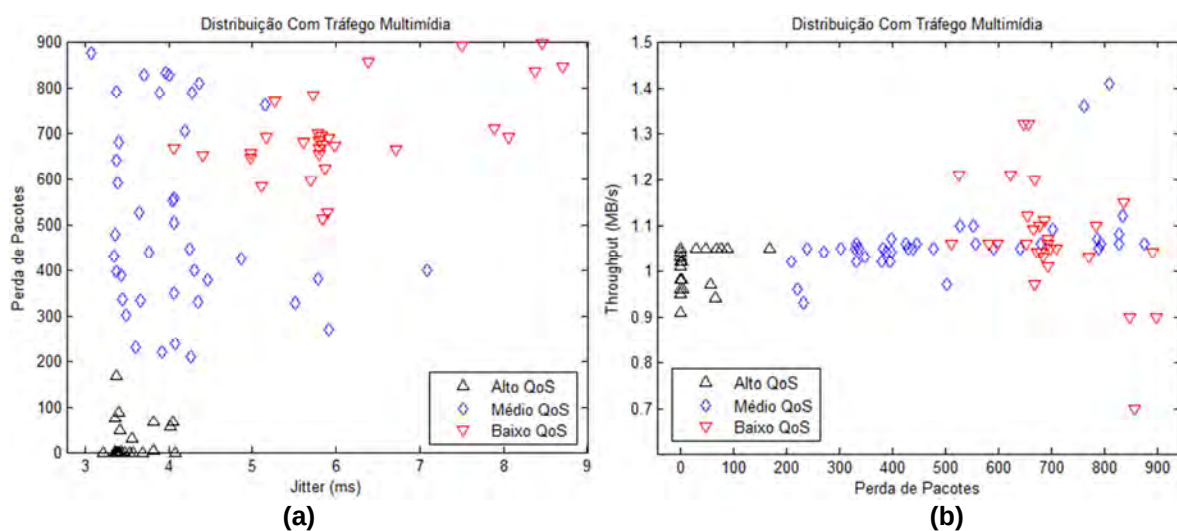


Figura 3. Rede MLP (RNA 3) com tráfego multimídia: (a) Melhor caso Jitter x Perda de Pacotes (b) Melhor caso Perda de Pacotes x Throughput.

5. Conclusões

Neste artigo é proposta uma análise de uma rede VPN, composta por um Servidor VPN, estações clientes VPN, parâmetros de QoS e redes neurais. Esta rede foi avaliada utilizando como base métricas de desempenho, tais como: *Throughput*, *Jitter* e Perda de Pacotes.

Foi utilizado no experimento uma das técnicas de redes neurais MLP (MultiLayer Perceptron), com isso foi possível prever o comportamento de uma rede VPN, através da classificação em Alto, Médio e Baixo QoS. A rede MLP apresentou um melhor caso para a RNA 2 sem tráfego multimídia apresentando 13,33 % de erro e um melhor caso da RNA 3 com tráfego multimídia apresentando 2,22 % de erro.

Desta forma, confirma-se a utilização da rede MLP para a classificação de QoS em conteúdo multimídia, apresentando bons resultados quando relacionados aos erros de dados de treinos e validação nas métricas de QoS para a rede VPN.

6. Referências

- Bayesware Discoverer (2000), Bayesware Limited. <http://bayesware-discoverer-trial-version.software.informer.com/> Acessado em: 15 de jan de 2013.
- Dias, M. M. ; Silva Filho, L. A. ; Betini, R. C. ; Ramos, E. M. L. S. (2008). “A Utilização de Software Livre na Análise de QoS em Redes IP Utilizando Mineração de Dados.” IX Workshop de Software Livre, Porto Alegre - RS.
- Fernandes, Anita Maria da Rocha. (2008) “Inteligência Artificial: Noções gerais” -3 imp. Florianópolis: Visual Books, 2008.
- Gomes, Rafael Lopes, Júnior, José Jailton, Júnior, Waldir Moreira, Abelém, Antônio Jorge (2009) “QoE and QoS in Wireless Mesh Networks” Federal University of Pará, Belém, Pará, Brazil and Instituto de Engenharia de Sistemas e Computadores (INESC) Porto, Portugal. IEEE Latin-American Conference on Communications (IEEE LatinCom 2009).
- Gómez, Gerardo, González, Javier Poncela, Aguayo-Torres, Mari Carmen, Muñoz, José Tomás Entrambasaguas (2010) “QoS Modeling for End-to-End Performance Evaluation over Networks with Wireless Access” EURASIP Journal on Wireless Communications and Networking Volume 2010.
- Guedes, Bruno Miguel Marcos (2009) “Application layer QoS support for video services”, Dissertação de Mestrado Bolonha em Engenharia de Redes de Comunicações. Instituto Superior Técnico, Universidade Técnica de Lisboa.
- Lage, Gustavo Vieira (2009) “Utilização de técnicas de Mineração de Dados para análise de marcações na Web.” Trabalho de Conclusão de Curso. Pontifícia Universidade Católica De Minas Gerais.
- Schroder, Carla (2008) “Measure Network Performance with iperf” <http://www.enterprisenetworkingplanet.com/netos/article.php/3657236/Measure-Network-Performance-with-iperf.htm>, Acessado em: 10 de Jan de 2013.
- Y. Bai, Chu Y., Ito e MR (2011) "Dynamic end-to-end QoS support for video over the Internet" AEU - International Journal of Electronics and Communications, In Press, Corrected Proof, Available online 20 January 2011.

A TI como uma Ferramenta de Gestão Estratégica nas Operações de Uma Empresa do Setor de Logística

Marcelo Martins Bassetto, Adriano César Santana

Escola de Engenharia Elétrica, Mecânica e de Computação (EMC) | Universidade Federal de Goiás (UFG) | Goiânia, GO | Brasil

marcelo_ufg@yahoo.com.br, adriano@emc.ufg.br

Abstract. *Nowadays, companies in the logistics industry must offer services more efficient, innovative, lower cost and time. With this, the business area of Information Technology (IT) has the role of providing the tools for the results to be achieved. This article presents the main concepts of applied IT in logistics, and a case study conducted in a large enterprise sector, where he showed a reduction of operating and administrative costs of approximately 10%, and a significant increase in customer satisfaction from 75% to 92%, and finally concludes that IT aligned with business objectives is critical to effective organizational management.*

Resumo. *A cada dia, as empresas do setor de logística precisam oferecer serviços cada vez mais eficientes, inovadores e de menor custo e tempo. Com isso, a área de negócio de Tecnologia da Informação (TI) tem o papel importante de disponibilizar as ferramentas para que os resultados sejam alcançados. Este artigo apresenta os principais conceitos de TI aplicados na área de logística, e um estudo de caso realizado em uma grande empresa do setor, onde mostrou-se uma redução de custos operacionais e administrativos de cerca de 10%, além de um significativo aumento na satisfação dos clientes de 75% para 92%, e por fim conclui-se que a TI alinhada aos objetivos da empresa é fundamental para uma gestão organizacional eficiente.*

1. Conceito de Logística

A palavra Logística tem sua origem ligada à área militar, em que o diferencial estratégico para se ganhar uma guerra, era a capacidade de deslocamento de mantimentos, munições, primeiro socorros e equipamentos, a tempo e a ordem às suas tropas nos campos de batalha. [Novaes, 2004]

Segundo Christopher (1999), o termo logística pode-se definir como o processo de se gerenciar estrategicamente a aquisição, movimentação e armazenagem de materiais, peças e produtos acabados através da organização e seus canais de marketing, de modo a maximizar as lucratividades presente e futura.

Neste mesmo sentido, Novaes (2004), diz que a logística é o processo de planejar, implementar e controlar de maneira eficiente o fluxo e a armazenagem de produtos, bem

como os serviços e informações associados, cobrindo desde o ponto de origem até o ponto de consumo, com o objetivo de atender aos requisitos do consumidor.

Desta forma, considerando essas duas visões, percebe-se que Christopher (1999), define a Logística de forma mais operacional, conceito esse já considerado antigo, pois em momento algum é considerado a visão das informações no processo. Enquanto que Novaes (2004) insere no contexto o gerenciamento da cadeia de suprimentos, considerando o monitoramento do processo para atender os requisitos do cliente.

2. Tipos de Logística

Para Rodrigues (2006), a logística pode ser dividida em três grandes categorias:

a) Transporte – formado pelo modal de transporte utilizado, pela forma de relacionamento entre tais modais, pelo elemento transportador e por suas instalações complementares. Sendo considerado um dos principais componentes da Logística, pois impacta diretamente na qualidade do serviço através do prazo de entrega, segurança e confiabilidade dos produtos.

b) Distribuição – processo envolvido desde a saída da mercadoria armazenada até a entrega ao cliente. Esta etapa do processo logístico está ligada diretamente na satisfação do cliente, pois tem o contato direto com o cliente final, que irá receber o produto adquirido e constatar se o prazo foi cumprido e se tudo foi entregue conforme solicitado.

c) Armazenagem – tem o objetivo de guardar certo produto por um determinado tempo, garantindo a segurança e integridade da mercadoria. Este processo considera estoque, espaço de armazenagem, *layout* do armazém e formas de movimentação.

3. Sistemas de TI para Controle de uma Empresa do Setor Logístico

Neste artigo, foi realizado um estudo de caso em uma grande empresa do setor logístico, a qual oferece serviços de transportes, armazenagem e distribuição.

Os principais sistemas implementados atualmente na gestão desta empresa são:

- ERP (*Enterprise Resource Planning*): software para controle da gestão administrativa financeira, fiscal, contábil e de compras;
- TMS (*Transport Management System*): sistema de gestão de transportes, que envolve desde a emissão dos conhecimentos de transportes até a fase de entrega do produto no cliente final;
- WMS (*Warehouse Management System*): controla o processo de armazenagem, desde a entrada do produto no armazém até a saída para entrega ao cliente;
- Rastreadores e Roteirizadores: sistemas utilizados para rastrear e acompanhar toda a carga durante o transporte, definir a melhor rota de entrega e oferecer informações aos portais disponibilizados aos clientes via internet, que podem pesquisar em que situação e onde se encontram suas mercadorias;

A completa integração dos sistemas mencionados anteriormente, se torna um fator crítico de sucesso e de fundamental importância, uma vez que a informação gerada sustenta os processos decisórios, dando mais competitividade, tanto para a empresa que oferece e presta o serviço logístico, quanto para seu cliente, que através de portais e

integrações com seu próprio sistema tenham informações em tempo real e de forma confiável. O aumento da velocidade do fluxo de informações geradas pela integração e disponibilização dos sistemas acima, impacta na redução de custos e maior velocidade de atendimento, gerando maior satisfação dos clientes.

4. Estudo de Caso

A empresa pesquisada neste estudo de caso oferece seus serviços à região Sudeste, Sul, Nordeste e Centro-Oeste, e atua no mercado de logística a quase 20 anos. Ela possui certificados de qualidade como ISO:9001, SASSMAQ e PSQ. Com isso, tem-se uma grande oportunidade de levantamento e validação da importância do uso eficiente da TI em todo o processo logístico e em suas diversas etapas de tomadas de decisão.

A equipe de TI com o apoio da alta direção da empresa de logística, decidiram melhorar os recursos internos e disponíveis para seus clientes, com isso adotaram como metodologia de estudo a pesquisa de mercado avaliando o que era utilizado por outras empresas do setor logístico, além de entrevistas com seus clientes para entender suas principais necessidades e sugestões de melhorias no processo.

A estrutura tecnológica e informacional é baseada na integração entre o ERP, WMS, TMS e o sistema do cliente, realizado por um EDI (Intercâmbio Eletrônico de Documentos). O uso de EDI para integração com o sistema do cliente foi definido pela empresa devido a solicitação do seu principal cliente, que é responsável por 60% do faturamento mensal, pois o mesmo já adotava essa prática com outros fornecedores.

A partir dessa integração, o processo se inicia na geração do pedido do cliente, que cria um ticket no sistema de gestão da logística, que aciona a ordem de coleta. Com isso, inicia-se no sistema de rastreamento dos veículos o acompanhamento em tempo real do estado da coleta, gerado pelo rastreador, o qual alimenta os sistemas e disponibiliza num portal via internet o status do pedido.

Esta integração gerou redução nos custos, eliminando a digitação dos pedidos no sistema, onde o próprio cliente importava o arquivo no portal e alimentava no sistema da empresa de logística, disparando o fluxo de trabalho para iniciar o processo da coleta.

Após a coleta ser realizada e encaminhada ao armazém, os produtos são identificados pelos coletores, através dos códigos de barras, que alimentam o WMS, analisando toda sua estrutura e endereçando da melhor forma possível as caixas e *palletes* para serem armazenados e ficarem disponíveis para venda, no estoque do cliente.

Uma nova etapa no processo é a venda destas mercadorias pelo cliente, que analisa o estoque disponível e gera os pedidos, que automaticamente envia para a equipe do armazém as ordens para montagem de cada pedido solicitado. As baixas de estoque são realizadas com a leitura dos códigos de barras, que prepara a carga para ser entregue, acionando agora o sistema TMS, onde é realizada toda a montagem da carga, geração do conhecimento de transportes e manifesto para liberação de entrega gerando a melhor rota através do sistema roteirizador integrado ao mesmo. Ao liberar o veículo para sair com a mercadoria, o sistema aciona novamente o rastreamento, que vai alimentando o portal com as localizações das mercadorias até a entrega final no cliente.

A integração entre o sistema TMS, o rastreamento e o portal, agrega um diferencial para a empresa analisada, pois automatiza boa parte do processo e deixa sempre disponível

ao cliente, através do portal, a situação de todos os seus pedidos, sua localização e em que fase se encontram dentro do processo. Assim, novamente temos redução nos custos operacionais, pois o cliente consulta seus pedidos ao invés de solicitar a empresa de logística o *status*, gerando maior satisfação do cliente e transparência nas operações.

Outro ponto importante desta integração com o cliente é a possibilidade do acesso ao portal por qualquer dispositivo móvel e com interação com a equipe de logística, que ficam online e podem auxiliar o cliente se necessário. Existe também o envio de mensagens automáticas pelos rastreadores, tanto para a logística quanto para seu cliente, informando sempre que uma carga for entregue através do sistema de rastreamento.

5. Conclusão

Com a integração e automatização implantadas nesta empresa, obteve-se uma redução de custos operacionais e administrativos de cerca de 10%, com a redução do quadro de funcionários do departamento de logística, redução de horas extras devido a agilidade do processo, menor utilização de telefonia, entre outros itens. Também notou-se um aumento de produtividade do departamento de logística, que analisam as informações de forma mais rápida e detalhada, tomando ações mais precisas e em tempo menor.

Anualmente a empresa envia uma Pesquisa de Satisfação para todos os seus clientes, e é notório o aumento das respostas positivas a cada ano com a implantação de melhorias tecnológicas em todo o processo logístico. Comparando-se o ano de 2008, ano em que a empresa não possuía integrações e alguns recursos citados neste estudo, com o ano de 2012, pode-se notar que a satisfação geral dos clientes aumentou de 75% para 92% conforme consolidação das respostas obtidas na pesquisa.

Com as informações apresentadas nota-se que a TI está presente em todo o processo logístico e é fundamental para o controle das informações envolvidas. Sendo que a TI deve oferecer as ferramentas mais adequadas e com o melhor nível de detalhe e disponibilidade, tanto para a equipe que faz a gestão da logística quanto ao cliente.

Conclui-se que a TI alinhada aos objetivos da empresa é fundamental e de extrema importância para a gestão organizacional, e que se aplicada de maneira correta é um diferencial competitivo, que auxilia nas tomadas de decisões gerando satisfação dos clientes, reduzindo os custos e consequentemente gerando lucros para as empresas.

6. Referências

- Christopher, Martin. (1999) "Logística e Gerenciamento da Cadeia de Suprimentos – Estratégias para redução de custos e Melhoria dos Serviços". 2ª edição. São Paulo, Pioneira.
- Novaes, Antônio G. (2004) "Logística e Gerenciamento da Cadeira de Distribuição". 3ª edição. Rio de Janeiro: Campus.
- Rezende, D. A. (2002) "Tecnologia da informação integrada à inteligência empresarial: alinhamento estratégico e análise da prática nas organizações". São Paulo: Atlas.
- Rodrigues, Paulo R. A. (2006) "Introdução aos Sistemas de Transportes no Brasil e à Logística Internacional". 3ª edição. São Paulo, Aduaneiras.

Os desafios para a mobilização de aplicações baseadas em plataforma Web

Celly de Siqueira Martins, André Lara Temple de Antonio, Carlos Antonio de Oliveira

Fundação CPqD – Centro de Pesquisa e Desenvolvimento em Telecomunicações
Campinas – SP – Brasil

Diretoria de Soluções em Billing

{celly, andret, coliveir}@cpqd.com.br

Abstract. *The use of mobile devices, such as tablets and smartphones, has significantly increased over the last years, and is changing the way people relate to one another and to their jobs. Wireless devices come in multiple shapes and sizes and such diversity in technology brings new challenges to the Information Technology sector, which are now addressed by the Mobile Computing Technology professionals. This paper describes the main challenges within the mobile computing context, types of mobile applications, in addition to two applications with different solutions, one of them migrated from a Web platform.*

Resumo. *O uso de dispositivos móveis, tais como tablets e smartphones, tem aumentado significativamente nos últimos anos e está transformando a forma de as pessoas se relacionarem com o trabalho e com os amigos. Esses dispositivos utilizam recursos de tecnologia sem fio e são apresentados em diversos formatos e tamanhos. A tecnologia de computação móvel surgiu para tratar essa diversidade tecnológica e traz novos desafios à área de tecnologia da informação. Neste artigo, são apresentados os principais desafios da computação móvel e são definidos os tipos de aplicativos móveis, bem como são demonstradas duas aplicações com distintas soluções, uma delas migrada a partir de uma plataforma Web.*

1. Introdução

O acesso à Internet a baixo custo a partir das redes sem fio, tais como Wi-Fi ou WiMAX, e o aumento significativo de dispositivos móveis com capacidade de processamento e comunicação estão provocando uma mudança na forma de desenvolver aplicações para esses dispositivos. Novos problemas e desafios surgiram para os profissionais da área de tecnologia de informação (TI) para tratar esse assunto denominado computação móvel.

Com o advento e a proliferação mundial de dispositivos móveis, a computação tornou-se verdadeiramente difundida [Figueiredo e Nakamura 2003; Reis, Carmo e Soto 2013]. A maior vantagem para os usuários de dispositivos móveis é acessar seus recursos e serviços sem a necessidade de manter-se conectado a uma infraestrutura fixa

e, em geral, estática. O acesso pode ser feito em qualquer lugar e a qualquer momento. A computação móvel possibilita aos usuários esse acesso independentemente de sua localização, podendo, inclusive, estar em movimento [Mateus e Loureiro 2013; Ogliari et al. 2006; Rishpater 2001].

2. Os desafios da computação móvel

Os aplicativos móveis fizeram surgir novos modelos de negócio, incluindo modelos centrados de dispositivos, em que o usuário pode ter acesso a novas aplicações e serviços, conectando-se a uma loja de aplicativos de fabricantes de dispositivos móveis, como, por exemplo, Google Play Store e Apple App Store. O principal fator dessa mudança é a capacidade avançada dos dispositivos móveis de utilizar as vantagens do processo de convergência, agregando serviços e aplicações avançadas de Internet ao próprio aparelho.

No entanto, o mercado de dispositivos móveis é dominado por diferentes plataformas tecnológicas, incluindo sistemas operacionais e plataformas de desenvolvimento de software, o que gera uma variedade de diferentes soluções disponíveis no mercado. Sendo assim, um dos principais desafios da computação móvel é o desenvolvimento de software em um ambiente de tecnologia altamente fragmentado e em rápida evolução. Os aplicativos móveis exigem uma razoável customização para que possam ser continuamente executados em diversas plataformas e versões.

Permitir que usuários acessem informações contextualmente relevantes a qualquer momento e em qualquer lugar é uma tarefa que apresenta desafios substanciais aos profissionais da área de TI. Os principais itens a serem tratados no âmbito de computação móvel estão listados a seguir [Figueiredo e Nakamura 2003; Hammershoj, Sapuppo e Tadayoni 2010; Kumkum 2010; Ogliari et al. 2006; Terry 2008]:

- Selecionar o tipo do aplicativo móvel – É preciso escolher entre o nativo e Web. Para isso, algumas questões devem ser respondidas. Entre elas: O conteúdo sofre muitas alterações? O aplicativo vai utilizar recursos do dispositivo móvel? Existe necessidade de divulgação do produto? É imprescindível um tempo rápido de resposta? A partir das respostas a essas perguntas, é preciso ponderar as vantagens e as desvantagens de cada tipo para então selecionar o mais apropriado ao contexto.
- Tratar conexões de rede sem fio – As conexões com rede sem fio costumam ser intermitentes, de baixa velocidade e com alta latência. O uso da funcionalidade assíncrona pode compensar o problema de desconexão, e o aplicativo pode continuar sendo executado até que a conexão seja retomada.
- Aprimorar a interface gráfica – A interface gráfica deve ser aprimorada para tratar interfaces em dispositivos menores, cuja área de visualização da maioria dos aplicativos é pequena. É preciso considerar a existência de diversos tamanhos e resoluções de telas.
- Tratar a entrada de dados – A entrada de dados nem sempre inclui um mouse; o recurso mais utilizado atualmente é o toque na tela.

- Considerar as características dos dispositivos móveis – Entre elas estão a utilização de recursos do aparelho, o armazenamento de dados e a dependência de baterias para poder funcionar.
- Considerar os riscos de segurança – A segurança abrange a autenticação de usuário, a integridade de dados e a prevenção a ataques maliciosos.

3. Aplicativos móveis

Selecionar o tipo de aplicativo móvel é um dos maiores desafios da computação móvel. Os tipos de aplicativos móveis são: aplicativos móveis Web e aplicativos móveis nativos [Castledine, Eftos e Wheeler 2011; Hammershoj, Sapuppo e Tadayoni 2010; Lionbridge 2012; Longoria 2004].

3.1. Aplicativos móveis Web

Um aplicativo móvel Web é uma aplicação Web formatada para dispositivos móveis. O usuário acessa esse aplicativo por meio de um dispositivo móvel, digitando uma URL em um navegador Web. Assim como um aplicativo Web tradicional, é construído com três tecnologias principais: HTML, CSS e JavaScript. No HTML são definidos os textos estáticos e as imagens; no CSS, o estilo e a apresentação e no JavaScript, as interações e animações. Uma vez que os aplicativos Web são baseados em navegadores Web, eles são independentes de plataforma e de dispositivos e podem ser executados em qualquer dispositivo móvel habilitado a acessar a Internet.

3.1.1. Vantagens dos aplicativos móveis Web

As principais vantagens dos aplicativos móveis Web são:

- Os aplicativos móveis Web são mais baratos e sua manutenção é mais simples. Isso ocorre pelo fato de disporem de compatibilidade multiplataforma e não precisarem lidar com as mudanças de diferentes dispositivos móveis.
- Facilidade de acesso – O usuário não precisa baixar uma aplicação, basta acessar uma URL por meio de um navegador Web e sempre obterá a aplicação mais atualizada para o dispositivo móvel.

3.1.2. Desvantagens dos aplicativos móveis Web

As principais desvantagens dos aplicativos móveis Web são:

- Apresentam tempo de resposta mais lento pelo fato de utilizarem HTML, serem baixado de um servidor Web a cada vez que é executado e consequentemente dependerem da conectividade e da velocidade da rede.
- Geralmente, não podem acessar as funcionalidades de hardware e de software dos dispositivos móveis. Se os requisitos exigirem, como, por exemplo, controle de câmera ou de GPS, integração com o gerenciador de informações pessoais ou controle de aplicativo do telefone, a opção por aplicativos móveis Web deve ser imediatamente descartada.

- Requerem uma conexão com a Internet para funcionar – Esse problema pode ser contornado com o uso da linguagem HTML5, a versão mais recente do HTML, uma vez que permite ao dispositivo móvel funcionar sem estar conectado.

3.2. Aplicativos móveis nativos

Um aplicativo móvel nativo é construído especificamente para determinado dispositivo móvel e seu respectivo sistema operacional. Ao contrário de um aplicativo Web, que é acessado por meio da Internet, um aplicativo nativo é baixado de uma loja de aplicativos e é instalado no dispositivo móvel.

Esses aplicativos são mais adequados para as aplicações que exigem frequentes atualizações, como, por exemplo, jogos, placares esportivos, relatórios de tráfego e sessões de bate-papo.

3.2.1. Vantagens dos aplicativos nativos

As principais vantagens dos aplicativos nativos são:

- Conteúdo mais atrativo ao usuário – Os aplicativos nativos podem utilizar os recursos nativos do dispositivo móvel, tanto de hardware (GPS, câmera, reconhecimento de voz e gráficos) como de software (e-mails, calendário, contatos, fotografias, vídeos e gerenciador de arquivos). Assim, o conteúdo pode ser mais atrativo e é possível tirar vantagem da capacidade do dispositivo móvel.
- Funcionalidade assíncrona – Como o aplicativo permanece instalado no dispositivo, ele pode ser executado mesmo que caia a conexão com a Internet.
- Rapidez no tempo de resposta – Os aplicativos móveis nativos são mais rápidos, pois, não precisam ser interpretados.
- Melhor divulgação do aplicativo aos usuários – Como os aplicativos móveis são a “atração do momento”, os usuários buscam incessantemente os aplicativos mais recentes e que despertam seu interesse. Isso oferece mais visibilidade aos produtos disponíveis nas lojas de aplicativos móveis e, conseqüentemente, maior probabilidade de serem descobertos.

3.2.2. Desvantagens dos aplicativos nativos

As principais desvantagens dos aplicativos nativos são:

- As informações de quem são os seus assinantes ficam públicas nas lojas de aplicativos.
- Para um aplicativo nativo ser executado em diversos dispositivos são necessárias diversas versões em diversas plataformas.
- Manter o aplicativo nativo atualizado também é trabalhoso, requer desenvolvimento, testes e distribuição para diferentes plataformas.

4. Soluções para aplicativos móveis

4.1. Aplicação CPqD liga+

O custo com ligações para telefones celulares de uma mesma operadora costuma ser mais barato ou até mesmo inexistente, dependendo do plano adquirido pelo usuário. Identificar antecipadamente a operadora de um número de telefone a ser chamado pode ajudar um usuário a melhor gerenciar suas ligações e economizar recursos financeiros.

A aplicação CPqD liga+ foi desenvolvida com o intuito de identificar a operadora de um número a ser chamado e de todos os números de contatos de uma agenda de um dispositivo móvel [Google Play 2013]. Além disso, é possível visualizar estatísticas de contatos da agenda agrupadas por operadora. A Figura 1 mostra uma imagem de telas da aplicação CPqD liga+.

A solução adotada para o desenvolvimento dessa aplicação foi baseada em aplicativos nativos. Essa escolha deu-se fundamentalmente pela necessidade de utilizar os recursos do próprio dispositivo. A obtenção dos números da agenda é feita por meio de uma integração com o gerenciador de informações pessoais. As chamadas utilizam o discador para serem efetuadas automaticamente.

Foi utilizada a plataforma de desenvolvimento Android e foram criados quatro layouts de telas para atender aos diferentes tamanhos (pequeno, normal, grande e extragrande) e às diferentes densidades (baixa, média, alta e muito alta) [Android 2013].



Figura 1. Imagem de telas da aplicação CPqD liga+

4.2. Aplicação CPqD Selfcare Mobile

CPqD Selfcare é um portal Web de autoatendimento voltado para o mercado corporativo e de varejo de telecom. Seu objetivo é fornecer ao cliente final uma interface amigável para consulta de faturas, controle de gastos, compra de produtos, serviços, solicitação e acompanhamento de solicitações, registro de reclamações ou elogios e demais funcionalidades que a operadora queira disponibilizar para seu cliente.

A aplicação CPqD Selfcare Mobile surgiu da necessidade de migrar algumas funcionalidades do mercado corporativo do produto CPqD Selfcare para dispositivos móveis. As principais funcionalidades são: consultar gráficos de consumo, consultar

faturas, apresentar mensagens e efetuar login. A Figura 2 exibe telas dessas principais funcionalidades.

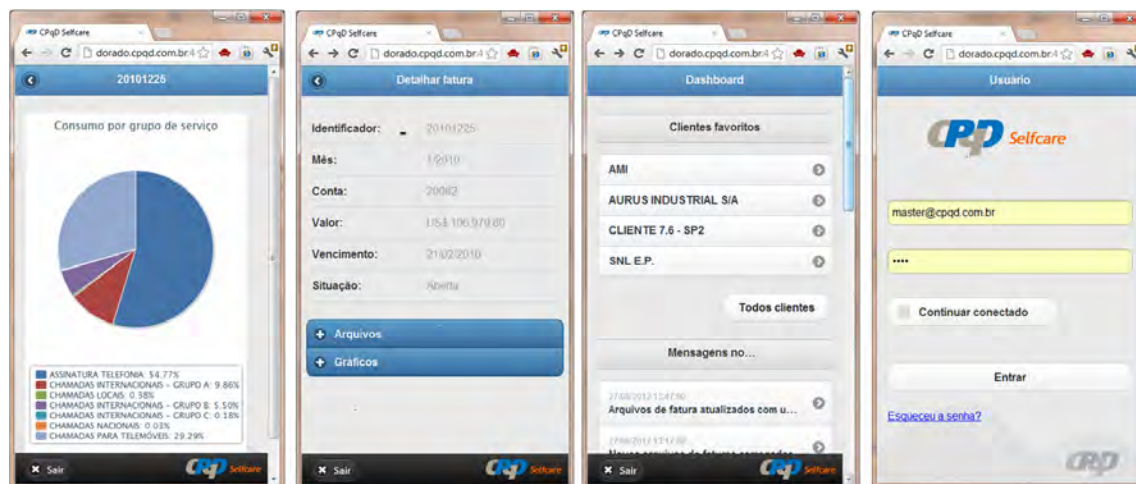


Figura 2. Imagem de telas da aplicação CPqD Selfcare Mobile

A solução adotada para o desenvolvimento dessa aplicação foi baseada em aplicativos Web. Essa escolha deu-se pelo fato de a aplicação não sofrer muitas alterações, o tempo de resposta não ser um requisito essencial e, principalmente, não precisar utilizar os recursos nativos dos dispositivos móveis.

5. Conclusão

A mobilização de aplicações baseadas em plataformas Web está em sua fase inicial e ainda são muitos os desafios existentes para os profissionais de TI. Definir o melhor tipo de aplicativo móvel a ser utilizado é o maior desafio. É preciso alinhar os requisitos funcionais da aplicação com as vantagens e as desvantagens de cada tipo de aplicativo e ponderar pela melhor solução.

Os aplicativos móveis Web costumam ser mais baratos por ter compatibilidade multiplataforma, mas não utilizam recursos dos dispositivos e podem apresentar lentidão no tempo de resposta. Os aplicativos móveis nativos costumam ser mais caros pelo fato de necessitar de especialistas para desenvolvimento e pela necessidade de se desenvolver um aplicativo diferente para cada plataforma. Entretanto, eles possuem uma grande vantagem no que diz respeito à possibilidade de utilizar os recursos nativos dos dispositivos móveis, tornando-se mais atrativos aos usuários e apresentando rápido tempo de resposta.

6. Trabalhos futuros

A tecnologia de computação móvel traz novos desafios aos profissionais da área de TI. Alguns trabalhos podem ser desenvolvidos para um melhor entendimento e aplicação da computação móvel. Entre eles é possível destacar:

- Estudo de HTML5 integrado a bibliotecas javascript, tais como Titanium, Rhodes e Phonegap, de forma a utilizar os recursos nativos dos dispositivos móveis, e, ainda assim, permitir a execução multiplataforma e usar a funcionalidade assíncrona para as quedas de conexões de redes sem fio.

- Desenvolvimento de novas aplicações em diversos domínios do conhecimento para explorar as possibilidades da computação móvel.

Referências

- Android (2013). “Grouping Resource Types”. Disponível em: <<http://developer.android.com/guide/topics/resources/providing-resources.html#AlternativeResources>>. Acesso em: jan. 2013.
- Castledine, E., Eftos, M. e Wheeler, M. (2011). “Build Mobile Websites and Apps for Smart Devices. Sitepoint.”
- Google Play (2013). “liga+ CPQD”. Disponível em: <<http://play.google.com/store/apps/details?id=com.cpqd.billing.operadora&feature=personalized-recommendations#?t=W251bGwsMSwyLDIxMywiY29tLmNwcWQuYmlsbGluZy5vcGVyYWRvcmlEiXQ>>. Acesso em: jan. 2013.
- Figueiredo, C. e Nakamura, E. (2003). “Computação Móvel: Novas Oportunidades e Novos Desafios”. T&C Amazônia, ano 1, n. 2. p. 16-28, Disponível em: <http://www.fucapi.br/tec/imagens/revistas/ed02_completo.pdf>. Acesso em: jun. 2012.
- Hammershoj, A., Sapuppo, A. e Tadayoni, R. (2010). “Challenges for mobile application development”. In: *Intelligence in Next Generation Networks (ICIN), 2010 14th International Conference on*, p. 1-8. IEEE.
- Kumkum, G. (2010). “Mobile Computing: Theory and Practice”. Pearson Education India.
- Lionbridge (2012). “Mobile Web Apps vs. Mobile Native Apps: How to Make the Right Choice”. Disponível em: <<http://en-us.lionbridge.com/kc/mobile-web-apps-vs-mobile-native-apps.htm>>. Acesso em: jan. 2013.
- Longoria, R. (2004). “Designing Software for the Mobile Context: A Practitioner's Guide”. Springer-Verlag. New York Incorporated. 153 p.
- Mateus, G. e Loureiro, A. (2013). “Introdução à Computação Móvel”. Disponível em: <http://homepages.dcc.ufmg.br/~loureiro/cm/docs/cm_livro_1e.pdf>. Acesso em: jan. 2013.
- Ogliari, R. et al. (2006). “Migrando Sistemas Legados de Tomada de Decisões para Ambientes Móveis”. In: *II Congresso Sulcatarinense de Computação*. Anais do II Congresso Sulcatarinense de Computação.
- Reis, C., Carmo, M. e Soto, C. (2013). “Computação Móvel: Tópicos Essenciais”. Disponível em: <<http://www.async.com.br/~kiko/mobilcomp/2.php>>. Acesso em: jan. 2013.
- Rishpater, R. (2001). “Desenvolvendo Wireless para a Web: Como enfrentar os desafios dos Projetos para a Web Sem Fio”. São Paulo, SP: Makron Books, 1ª. edição. 350 p.
- Terry, D. (2008). “Replicated Data Management for Mobile Computing”. *Synthesis Lectures on Mobile and Pervasive Computing* 3, n. 1: p.1-94.

Lógica para Semáforo Inteligente Baseado na Mineração de Dados por Algoritmo Genético Transgênico

Lucas Costa Cunha¹, Luciana Aparecida Ferreira¹, Fabrício Alves Rodrigues²,
Wanderley de Souza Alencar¹

¹Curso de Ciência da Computação - Universidade Federal de Goiás/Jataí (UFG)
BR364 Km 192, Setor Industrial, Jataí - GO - Brasil

²Pós-Graduação em Ciência da Computação - Universidade Federal de São Carlos (UFSCar)
Rodovia Washington Luís, km 235 - SP-310 São Carlos - São Paulo - Brasil

l.cunha@hotmail.com, fabricio1989@gmail.com

Abstract. *In large and medium size cities there is a known problem by traffic engineers, traffic flow problem. In computer science this problem is seen in the area of computer networks when there is a congestion somewhere in the network. This article proposes a solution for traffic flow problem using heuristics of Artificial Intelligence, Genetic Algorithms, Data Mining to make a pattern and get specific rules which will be used on the rules fuzzy, for the automation of a traffic light to traffic flow at a certain passage, when there is a need for change to see as this environment.*

Resumo. *Em grandes e medias cidades há um problema já conhecido pelos engenheiros de tráfego, o problema de fluxo de tráfego. Em ciência da computação esse problema é visto na área de redes de computadores quando existe um congestionamento em algum ponto da rede. Este artigo propõe uma solução para o problema de fluxo de tráfego utilizando de heurísticas da inteligência Artificial, Algoritmos Genéticos, Mineração de Dados para criar um padrão e estabelecer regras específicas que serão utilizadas nas regras Fuzzy, para a automação de um Semáforo para que o tráfego flua, em determinado trecho, quando existir a necessidade da mudança ao ver como está o ambiente.*

1. Introdução

Os esforços para a modelagem de problemas de ET (Engenharia de Trânsito) tem se tornado uma alternativa promissora. Problemas de ET, um dos problemas é o fluxo de tráfego, são formulados como problemas de programação linear inteira, ou seja, se tem uma função de otimização e suas restrições para serem analisadas é do tipo NP-completo. A solução destes problemas pode ser obtida em tempo hábil, através de métodos aproximativos [Dias 2004].

Para resolver o problema de fluxo de tráfego muitos pesquisadores tem estudado formas de mudar o tempo do semáforo e assim melhorar um pouco o problema do fluxo em um determinado trecho da malha de tráfego urbano, nos últimos anos tem se tornado um grande desafio a otimização dos tempos do semáforo [Lämmer and Helbing 2008]. O trabalho de Tonguz et al. [Tonguz et al. 2009] apresentou uma nova abordagem de autômatos celulares para a construção de um modelo de tráfego urbano e assim encontrou algumas características dos padrões globais de tráfego nas áreas estudadas e como

resultados mostrou que controles como a duração do ciclo, *split* verde, e coordenação de semáforos mostraram bons resultados para a dinâmica de tráfego.

Já nos trabalhos [Karakuzu and Demirci 2010, Yang et al. 2012, Askerzade and Mahmood 2010] é proposto um semáforo com um controle inteligente para mudanças de fases do sinal, vermelho e verde, utilizando lógica *fuzzy* com objetivo de diminuir o problema do fluxo de tráfego. Para validar os resultados foi desenvolvido um simulador para representar o problema de fluxo com semáforos sincronizados e usando a lógica *fuzzy*.

Com o aumento significativo do uso dos automóveis e a expansão do sistema de trânsito surgiu o estudo de como encontrar características do tráfego, envolvendo modelos matemáticos [Prado 2010]. Neste trabalho será utilizado algoritmo genético, que é um método heurístico que busca os melhores resultados criado por Holland que faz uso do processo evolutivo descrito por Darwin, para descrever uma característica, ou também um padrão, para o tráfego.

O Algoritmo Genético pode ser utilizado para fazer a mineração dos dados como mostra os trabalhos [Amaral 2007, Bevilaqua et al. 2011, Minaei-Bidgoli and Punch 2003, Shah and Kusiak 2007, Cox 2005]. A mineração de dados foi desenvolvida na década de 80 para encontrar algum tipo de padrão em grandes base de dados e é muito utilizado também para geração de novos conhecimentos, neste trabalho os resultados da mineração de dados serão utilizados na lógica *Fuzzy*, que é uma lógica criada por Zadeh que utiliza grau de pertinência, entre 0-1, dos dados analisados e desta forma criar regras difusas para fazer inferência para alcançar a solução do problema.

O Objetivo deste trabalho é mostrar a eficiência do Algoritmo Genético Transgênico, descrito por [Amaral and Jr. 2011], para minerar dados, pois na bibliografia pesquisada não foi encontrado nenhum trabalho que propôs minerar uma base de dados de tráfego para ser criar uma regra difusa, também mostrar a eficiência da lógica difusa para realizar controle de sinal do semáforo e para diferenciar de outros trabalhos será utilizado a mineração de dados em uma base de dados que representa o fluxo de tráfego de um trecho de Jataí, para que a regra *fuzzy* realmente represente o comportamento do tráfego.

Para mostrarmos a eficiência da lógica *Fuzzy* será utilizado o seguinte estudo de caso: Em um trecho de trânsito T1 há um Semáforo que coordena o tráfego daquele trecho, pesquisadores da engenharia de tráfego analisaram o fluxo do trecho T1 por três anos, levando em conta a data, os horários, se houve ocorrência de chuva, quantidade de fluxo naquele período de tempo e se houve ou não congestionamento.

Este estudo de caso tem como base dados recolhidos na cidade de Jataí no trecho da avenida Goiás com a avenida Brasil e será utilizado para mostrar que os métodos de IA (Inteligência Artificial) e as regras difusa podem ajudar o semáforo a tomar uma decisão sobre como irá funcionar o fluxo naquele trecho, levando em conta a hora e a quantidade de fluxo naquele trecho estudado.

2. Referencial Teórico

O estudo científico do fluxo de tráfego teve seu início na década de trinta aplicando a teoria de probabilidade para descrever o comportamento do tráfego na estrada e com modelos

de estudos que relacionavam o volume, a velocidade e a investigação do desempenho do tráfego nos cruzamentos [Prado 2010].

O fluxo do tráfego, com observação de dados reais, é classificado como sendo um efeito espaço-temporal: existe a ocorrência de congestionamento no espaço e no tempo dentro de uma rede de tráfego. Estes padrões empíricos de via congestionada exibem algumas características complexas. Por esta razão, a complexidade da gestão do tráfego está associada com esta variedade de seus padrões, bem como com a necessidade na otimização. Esta otimização pode resolver o problema de congestionamento ou, se isto não é possível alcançar, a minimização da influência de uma situação de via congestionada [Kerner 2009].

Para manipular dados complexos que podem estar associados a um padrão é utilizada a mineração de dados. Em meados de 1980, a Mineração de Dados (MD) (ou também chamada de *Data mining*) foi criada por diversos pesquisadores e grandes empresas que tinham como objetivo criar métodos que conseguissem resolver os desafios descritos acima e para analisar os dados que estavam sendo gerados.

A MD se baseia em abordagens como amostragem, estimativa, teste de hipóteses, algoritmos de busca, teorias da inteligência artificial, reconhecimento de padrões e aprendizado de máquina. A MD atraiu vários pesquisadores de áreas diversas todos com o mesmo objetivo de buscar bons resultados a fim de lidar com o crescimento dos bancos de dados e com sua alta dimensionalidade [Tan et al. 2009].

A mineração de dados é referente a extração de conhecimento dentro de grandes conjuntos de dados, também é associada ao KDD (*Knowledge Discovery Database*), a MD apresenta ao usuário padrões encontrados na base de dados, que muitas das vezes o usuário não encontraria facilmente [Dantas et al. 2008].

Existem vários métodos para fazer a mineração de dados, no *software* Weka [Homes et al. 1994] pode-se encontrar muitos tipos de mineradores, neste artigo será utilizado o algoritmo genético (AG), como minerador de dados.

Os Algoritmos Genéticos são algoritmos de busca, que utilizam o processo da seleção natural proposto por Charles Darwin, os AGs foram inicialmente propostos por John Holland [Holland 1975]. Os AGs são bastante utilizados em problemas que dado um conjunto de elementos ou indivíduos se alcance aquele ou aqueles, dependendo do problema proposto, que melhor atendam a certas condições previamente especificadas [Silva 2001]. Os AGs são recomendados para problemas NP-Completo e NP-Difíceis, devido ao tempo que AG leva para processar os dados [Linden 2008].

Para fazer a busca do melhor resultado o AG executa os seguintes operadores:

- Seleção: Um dos meios de fazer a seleção de indivíduos para o *crossover*, e o mais simples, é o método da roleta no qual cada indivíduo da população recebe uma fatia da roleta, essa fatia é determinada pela avaliação (*Fitness*) do indivíduo desta forma quem estiver mais próximo da solução do problema terá mais chances de ser sorteado. Em seguida é feito o sorteio aleatório, seria análogo ao "rodar a roleta", desses indivíduos os selecionados farão parte do operador de *crossover* e são colocados em uma população auxiliar [Man et al. 2001].
- Cruzamento ou *Crossover*: O operador de cruzamento e de mutação são operadores chave do AG para que a busca seja bem sucedida, são os operadores mais

importantes do AG [Man et al. 2001]. Existem várias estratégias para esta operação será utilizada neste artigo a estratégia de cruzamento de um ponto, um ponto de cruzamento é escolhido aleatoriamente e a partir dele é feita a troca genética dos indivíduos escolhidos na seleção (pais) para a geração dos novos indivíduos (filhos). As informações genéticas anteriores a este ponto de um dos pais é ligado as informações posteriores do outro pai gerando a cada cruzamento dois novos filhos, que são colocados na população auxiliar [Rezende 2005].

- **Mutação:** O operador mutação é necessário para que a população seja diversificada, permite chegar mais rápido ao objetivo da busca e para fugir dos mínimos locais, alterando aleatoriamente um ou mais genes dos indivíduos gerados pelo cruzamento.
- **Reinserção:** Para saber quais indivíduos da população auxiliar será colocado na nova população é utilizado o operador de reinserção. Com a estratégia do elitismo pode haver um crescimento grande de super indivíduo na população, mas usando uma estratégia mais equilibrada parece melhorar o desempenho do AG, essa estratégia é chamada de melhores pais e filhos onde são escolhidos $N/2$ melhores indivíduos geradores (pais) e $N/2$ melhores indivíduos criados (Filhos) no qual são inseridos na população gerando uma nova geração [Man et al. 2001].

Em 2011 foi proposto pelos autores [Amaral and Jr. 2011] um algoritmo genético com um operador transgênico, esse novo operador para o algoritmo genético foi inspirado pela engenharia genética, onde permite manipular o material genético de indivíduos adicionando características que acredita-se ser a mais importante. Essa mesma atividade que é feita por engenheiros genéticos foi introduzido no AG, assim identificasse qual é o gene ou genes mais importante então esses genes dos N melhores indivíduos são passados para M outros indivíduos da população.

Um algoritmo genético, se utilizado para minerar dados, pode gerar regras e serem utilizadas em algum tipo de lógica. Lógica *Fuzzy* foi criada pois, havia algumas questões que a lógica booleana não conseguia afirmar com verdadeiro ou falso, então utilizando um grau de certeza que vai de 0-1 (0-100%) pode [Weber and Klein 2003].

O conceito *fuzzy* foi apresentado, em 1965, por Lotfi A. Zadeh [Zadeh 1965]; foi para ele o reconhecimento como colaborador dessa lógica e da inteligência artificial. porém o filósofo Platão já havia feito um estudo com uma terceira região de certeza, além do verdadeiro e falso [Weber and Klein 2003].

Em meados de 1960, Zadeh viu que recursos de automatização, de natureza industrial, biológica ou química, eram incapazes de compreender algumas situações ambíguas. Assim em 1965 o professor Zadeh publicou um artigo com os conceitos dos conjuntos *fuzzy* [Zadeh 1965] [Weber and Klein 2003].

A lógica *fuzzy* utiliza preposições linguísticas, exemplo Alto(x), Quente(x), Muito Rápido(x). Cada termo linguístico tem um valor de pertinência entre [0-1] [Ross 2010].

Existem três etapas na lógica *fuzzy* a primeira é a "fuzzificação", nessa etapa a aplicação difusa extrai os dados do mundo real e os transforma, por meio de uma função, em dados do universo difuso entre [0-1]. A segunda etapa é a inferência *fuzzy*, a aplicação difusa aplica as regras nos dados de entradas, já no universo [0-1], dado pela primeira etapa, e obtêm uma resposta. A terceira etapa é "defuzzificação", nessa etapa a aplicação

pega o resultado obtido pela regra que está no universo [0-1] e o transforma em uma saída com dados do mundo real por alguma função de "defuzzificação" [Ross 2010].

3. Solução Proposta

Para buscar um resultado favorável para o problema descrito neste artigo será utilizado algoritmo genético para minerar os dados e o conjunto *fuzzy* será usado para fazer a inferência do ambiente de tráfego do trecho, descrito na seção 1.

Será utilizado para fazer a mineração de dados o API Java do AG transgênico (AGT), criado pela Universidade Federal de Goiás *Campus* Jatai. Essa API (*Application Programming Interface*) segue os princípios descritos por [Holland 1975] e com o operador transgênico descrito por [Amaral and Jr. 2011]. O diagrama de atividades, figura 1, descreve como é o fluxo normal da API mostrada neste parágrafo.

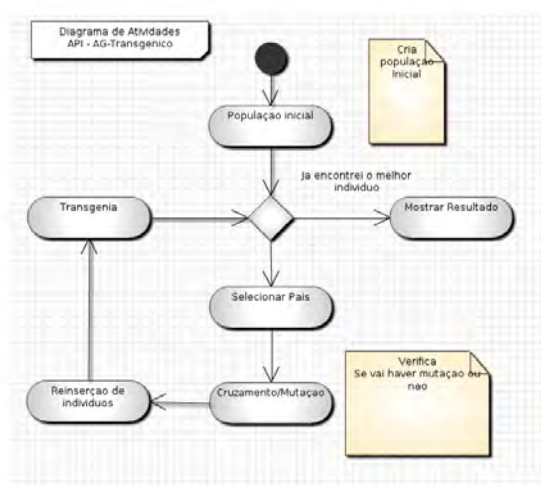


Figura 1. Diagrama de Atividades API AG-Transgênico

A configuração do cromossomo para o indivíduo do AG irá conter os genes: data (dia-mês), chuva com valores [Sim-Não], fluxo do tráfego valor inteiro e o congestionado, que será a classe para o minerador de dados, com valores [Sim-Não], o *fitness* ou avaliação do indivíduo se dá pela equação 1, os valores TN (*True Positive*, o minerador disse que era da classe 'Sim' e realmente era), FP (*False Positive*, o minerador disse que a classe não era 'Sim' e realmente não era) e os valores 0.5 relativos a 50% relativos aos dados [Sim,Não]. A possibilidade de existir uma mutação em algum indivíduo do AG será de 0,2%. O valor a ser buscado pelo AG é de um indivíduo com uma avaliação de 100%.

$$Fitness = TP * 0.5 + FP * 0.5; \quad (1)$$

A base de dados contém 25776 registros, que são os dados simulados que representa os três anos de coleta de dados dos pesquisadores como descrito na seção 1, a base apresenta uma quintupla que são os dados utilizados no cromossomo do indivíduo.

As regras difusas serão criadas de acordo com o resultado da mineração de dados levando em consideração o período e o fluxo do tráfego do trecho analisado e como termo lógicos serão:

Tabela 1. Resultados obtidos

Quantidade de testes	Itens Avaliados	AGT	AG
		Média	Média
		164	1250
		100%	100%
50	Gerações	43,553	87,341
	Fitness		
	Tempo de Execução(Segundos)		

Tabela 2. Linguísticas Para a Fuzzy

Conjuntos	valores de X[0-23]	valor de pertinência	Linguística
Período	$23 \geq X < 6$	$Y=0$	Baixo
	$6 \geq X < 12$	$Y=(N^{\circ} \text{ de registro do período } X \text{ na base de dados } * 1)/1800$	Médio
	$13 > X < 17$		
	$18 > X < 23$		
	$12 \geq X \leq 13$	$Y=1$	Pico
	$17 \geq X \leq 18$		
Fluxo	$0 \geq X \leq 20$	$Y=(X*1)/51$	Baixo
	$20 > X < 51$		Médio
	$X \leq 51$		Alto

- Período: Baixo, Médio e Pico;
- Fluxo: Baixo, Médio e Alto.

Logo após a criação das regras será usado um *software* para simular como será a reação do semáforo inteligente com a inferência *fuzzy*, o *software* irá obter as condições do trecho (Período e Fluxo) e com a ajuda das regras *fuzzy* tomar uma decisão.

4. Resultados Obtidos

Para fazer uma análise de como o AG transgênico pode ser utilizado para encontrar padrões em grandes *data sets* foi executado 50 vezes, como mostra a Tabela 1. Em todos os testes o minerador de dados com AG chegou ao resultado de que o congestionamento ocorre quando o tráfego está acima ou igual a cinquenta e um. É perceptível que o algoritmo genético transgênico teve grandes resultados se comparado com o algoritmo genético tradicional, esse resultado pode ser justificado pelo fato do transgênico guiar a população do AG a um possível ótimo.

Com os resultados trazidos pelo AG transgênico foi possível criar as fórmulas das linguísticas para a inferência *fuzzy*, essas fórmulas foram criadas pelos autores ao analisar a saída do minerador de dados, como mostra a Tabela 2. A saída da MD foi uma única regra que informava que classe congestionamento era sim se o valor de tráfego fosse ≥ 51 e se o valor fosse < 51 a classe seria não. As regras ficaram sendo as mostradas na Tabela 3, foi criada baseada na base de dados e na saída da mineração de dados, onde y é o grau de pertinência da entrada e s é o grau de pertinência da saída.

Foram testadas todas as possibilidades de ambiente e o simulador do semáforo respondeu corretamente, com as regras difusas, tendo assim uma resposta 100% para as inferências propostas neste artigo.

Tabela 3. Regras Nebulosas

Regras			
Rx	Período	Fluxo	Tempo Semáforo verde
R01	Medio(y) ou Pico(y)	Alto(y)	Alto(s)
R02	Baixo(y) ou Medio(y)	Medio(y)	Medio(s)
R03	Medio(y)	Baixo(y)	Medio(s)
R04	Baixo(y)	Baixo(y)	Baixo(s)

5. Conclusão e Trabalhos Futuros

Os resultados apresentados na seção anterior mostra que o AG Transgênico teve um desempenho de 50% melhor que o AG tradicional e o conjunto *Fuzzy*, baseado na mineração de dados, pode ser aplicado nos Semáforos de forma inteligente, onde o Semáforo toma sua decisão de acordo com o ambiente. Em todos os testes aplicados no simulador o Semáforo respondeu favorável o que demonstra que o objetivo do artigo foi válida.

Porém os testes foram feitos com um simulador utilizando dados reais e não foram feitos testes reais com o Semáforo, pois este artigo tem como objetivo mostrar a eficiência das inferências *fuzzy* com mineração de dados dentro de um problema da engenharia de tráfego. Para trabalho futuros pode ser feito uma melhor análise das regras. Utilizar mais de um Semáforo no teste para ver como seria a sincronização dos mesmos, será feita uma nova coleta de dados com novas variáveis para ser minerado e também, se possível, implementar em um semáforo real para ver como ele se comporta na prática.

Referências

- Amaral, L. R. (2007). Mineração de regras para classificação de oncogenes medidos por microarray utilizando algoritmos genéticos. Universidade Federal de Uberlândia. Dissertação de Mestrado.
- Amaral, L. R. and Jr., E. R. H. (2011). Transgenic, an operator for evolutionary algorithms. In *IEEE Congress on Evolutionary Computation*, pages 1308–1314. IEEE.
- Askerzade, I. N. and Mahmood, M. (2010). Control the extension time of traffic light in single junction by using fuzzy logic. *International Journal of Electrical e Computer Sciences*, 10(02):52–59.
- Bevilaqua, A., Rodrigues, F. A., and Amaral, L. R. (2011). Gasnp classifier: A machine learning environment for building high-level biological knowledge. *International Journal of Information Technology and Intelligent Computing*.
- Cox, E. (2005). *Fuzzy Modeling and Genetic Algorithms for Data Mining and Exploration*. The Morgan Kaufmann series in data management systems. Elsevier Science, USA, 1ª edição.
- Dantas, E. R. G., Junior, J. C. A. P., Lima, D. G., and Azevedo, R. R. (2008). O uso da descoberta de conhecimento em base de dados para apoiar a tomada de decisões. V *Simpósio de Excelência em Gestão e Tecnologia - SEGeT*, 01:50–60.
- Dias, R. A. (2004). Engenharia de tráfego em redes ip sobre tecnologia mpls: Otimização baseada em heurísticas. Universidade Federal de Santa Catarina, UFSC, Brasil. Tese de Doutorado.

- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, USA.
- Homes, G., Donkin, A., and Witten, I. H. (1994). Weka: A machine learning workbench. *Proceedings of the Second Australia and New Zealand Conference on Intelligent Information Systems*.
- Karakuzu, C. and Demirci, O. (2010). Fuzzy logic based smart traffic light simulator design and hardware implementation. *Appl. Soft Comput.*, 10(1):66–73.
- Kerner, B. S. (2009). *introduction to modern traffic flow theory and control the long road to three phase traffic theory*. Springer, Berlin, Alemanha.
- Lämmer, S. and Helbing, D. (2008). Self-control of traffic lights and vehicle flows in urban road networks. 0802.0403.
- Linden, R. (2008). *Algoritmos Genéticos: uma importante ferramenta da inteligencia computacional*, volume 3. Brasport Livros e Multimídia, São Paulo, Brasil.
- Man, K. F., Tang, K. S., and Kwong, S. (2001). *Genetic Algorithms: Concepts and Designs*. Spring, 3 edição.
- Minaei-Bidgoli, B. and Punch, W. F. (2003). Using genetic algorithms for data mining optimization in an educational web-based system. In *Proceedings of the 2003 international conference on Genetic and evolutionary computation: PartII*, GECCO'03, pages 2252–2263, Berlin, Heidelberg. Springer-Verlag.
- Prado, A. L. (2010). Modelagem matemática do fluxo de tráfego veicular. Universidade Federal do Paraná, UFPR, Brasil. Tese de Doutorado.
- Rezende, S. O. (2005). *Sistemas inteligentes: fundamentos e aplicações*. Manole Ltda, Barueri,SP.
- Ross, T. (2010). *Fuzzy logic with engineering applications*. John Wiley e Sons, Inglaterra, 3 edição.
- Shah, S. and Kusiak, A. (2007). Cancer gene search with data-mining and genetic algorithms. *Comput. Biol. Med.*, 37(2):251–261.
- Silva, E. E. (2001). Otimização de estruturas de concreto armado utilizando algoritmos genético. Escola Politécnica da Universidade de São Paulo. Dissertação de Mestrado.
- Tan, P.-N., Steinbach, M., and Kumar, V. (2009). *Introdução ao datamining*. Ciência Moderna, Rio de Janeiro, Brasil.
- Tonguz, O., Viriyasitavat, W., and Bai, F. (2009). Modeling urban traffic: A cellular automata approach. *Communications Magazine, IEEE*, 47(5):142 –150.
- Weber, L. and Klein, P. A. T. (2003). *Aplicação da Lógica Fuzzy em Software e Hardware*. Ulbra, Rio Grande do Sul, Brasil.
- Yang, W., Zhang, L., He, Z., Yang, Y., and Fang, Y. (2012). Urban traffic signal two-stage combination fuzzy control and paramics simulation. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 771 –775.
- Zadeh, L. (1965). Fuzzy sets. *Information and Control*, 8(3):338 – 353.

Motor de inferências em plataforma OSGi

Ding Yih An, Juan X. E. A. Calles, Karin Komati, Maxwell E. Monteiro

Coordenadoria de Informática – Instituto Federal do Espírito Santo (IFES) Campus Serra, Rodovia ES-010, Km 6.5 – 29.164-231 – Serra – ES – Brasil

{callesjuan,sidnha}@gmail.com; {kkomati,maxmonte}@ifes.edu.br

Abstract. *Most of the Ontology-based applications developed nowadays use OWL files for representing their ontology models, also they embed reasoners so that they can validate these OWL files. This work intends to show the benefits of providing reasoning as a service, also, briefly presents the steps for an implementation.*

Resumo. *A maioria das aplicações de base ontológica desenvolvida nos dias de hoje, usam arquivos OWL para representar seus modelos de ontologias. Elas também incorporam raciocinadores para que possam validar esses arquivos OWL. Este trabalho pretende mostrar os benefícios da prestação de raciocínio como um serviço, também apresenta brevemente os passos para a sua implementação.*

1. Introdução

Em definição curta, uma ontologia é uma especificação de uma conceituação (Gruber, 1993) que expressa os conceitos e as relações de um determinado domínio de forma não ambígua. Geralmente são representadas computacionalmente em formato OWL (*Web Ontology Language*), o que permite a definição de regras baseadas em lógica descritiva, chamadas de axiomas (Lindörfer, 2010). É possível fazer validações semânticas em arquivos OWL através de ferramentas chamadas de raciocinadores (*reasoners*), tais como: verificar a consistência da ontologia e inferir novas informações a partir da mesma. Os *reasoners* são utilizados em momentos específicos do ciclo de vida da aplicação: início de operação e nas atualizações das ontologias utilizadas.

Geralmente, o ferramental para a construção de aplicações baseadas em ontologias não suporta, diretamente, o paradigma de computação como serviço (Hewitt, 2009). Nesse contexto de computação como serviço, vale destacar o OSGi (*Open Services Gateway Initiative*), que é um conjunto de especificações que definem um sistema de componentes dinâmicos para Java, levando em consideração questões de empacotamento, versionamento, implantação, publicação e descoberta de serviços. O conceito fundamental que permite o funcionamento de um sistema desse tipo é a modularidade e o uso de interfaces de acesso a serviços. Módulos são componentes menores e reutilizáveis, aqui denominados de *bundles*, que escondem as suas implementações de outros componentes durante a comunicação por meio de serviços. Esses *bundles* são capazes de publicar serviços dinamicamente, além de descobrir e requisitar serviços disponibilizados por outros *bundles*. Uma característica importante desta especificação é que qualquer *bundle* pode ser executado em qualquer implementação do OSGi. (OSGi Alliance, 2012).

De forma geral, sistemas baseados em ontologias tem o *reasoner* embutido em si, o que aumenta a sua complexidade e a quantidade de recursos necessários para sua

execução. Em uma plataforma OSGi, por exemplo, estas questões se tornam evidentes quando se imagina que cada *bundle* envolvido em uma aplicação de ontologia deveria incluir em si uma instância distinta de um determinado *reasoner*.

O presente trabalho propõe a criação de um motor de inferências que visa atender o crescimento de sistemas distribuídos que fazem uso de ontologias, minimizando o trabalho de se ter que carregar múltiplas instâncias do *reasoner* dentro de seus módulos, já que o uso desses *reasoners* é limitado a certos momentos da sua operação. Desta forma, promove-se o reuso e diminui-se a complexidade da gestão de instâncias e artefatos relacionados, aproveitando as facilidades que uma plataforma OSGi oferece (Monteiro, 2010).

2. Solução Proposta

O *reasoner* escolhido para implementação deste trabalho foi o Pellet (Clark & Parsia, 2011) por ser um software consolidado e o primeiro a suportar inferências em artefatos OWL. A implementação OSGi escolhida foi o Eclipse Equinox (The Eclipse Foundation, 2013), por ser a implementação de referência da OSGi Alliance, a partir da revisão 4 da especificação OSGi. A utilização de outro *reasoner*, ou implementação OSGi, não deve afetar o resultado da solução proposta.

Para demonstrar o fornecimento de serviços semânticos na verificação de consistência e inferência de novas informações, foi criado um *bundle* OSGi que disponibiliza as funcionalidades do motor de inferência Pellet. Os serviços de raciocínio relevantes identificados na implementação são os de:

- consistência - certifica que a ontologia não contém quaisquer fatos contraditórios;
- classificação - calcula as relações de cada subclasse com as outras classes para descobrir todas as heranças existentes e criar uma hierarquia completa a partir disso e;
- realização - encontra a classe específica à qual cada um dos indivíduos pertence.

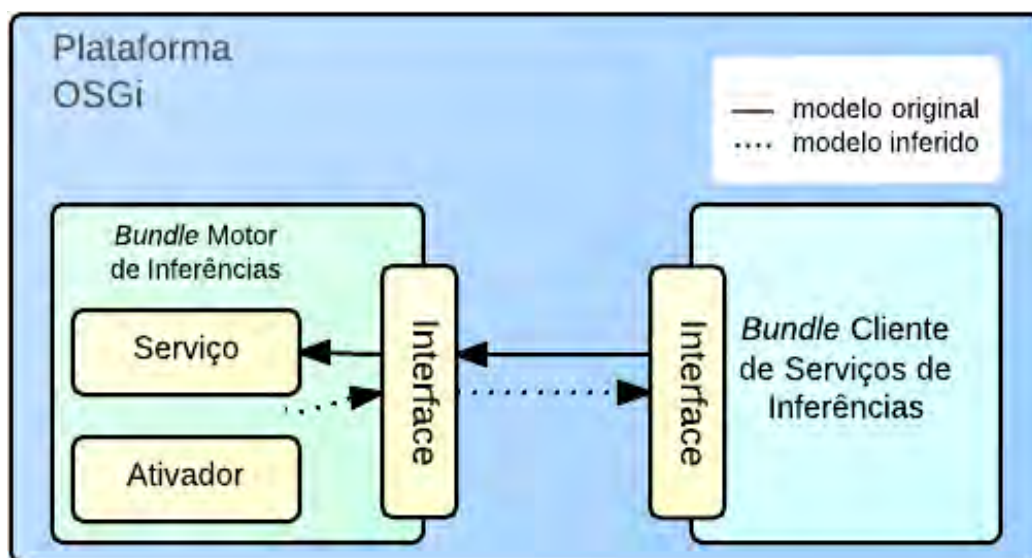


Figura 1 - Cliente acessando os serviços do *bundle* Motor de Inferências.

O modelo da solução é constituída por três partes, conforme visto na Figura 1:

- interface - que descreve os serviços oferecidos com os seus respectivos parâmetros;
- serviço - implementa os serviços de raciocínio descritos previamente na interface e;
- ativador - responsável por registrar e disponibilizar o serviço implementado, ou seja, pelo ciclo de vida do *bundle*.

O funcionamento se dá pela criação do *bundle* descrito acima, composto por interface, serviço e ativador, e em seguida especificam-se no arquivo de manifesto as interfaces dos próprios serviços que ele exportará. Finalmente, ele é registrado na plataforma através do ativador. As informações contidas no arquivo de manifesto serão disponibilizadas para os demais *bundles*. Para o consumo dos serviços, é preciso que clientes importem em seus arquivos de manifesto as interfaces do motor de inferências Pellet e referenciem os serviços no seu próprio ativador.

A partir da arquitetura definida, o cliente é capaz de utilizar os serviços citados anteriormente, fornecendo um modelo ontológico (modelo original) como entrada (como mostrado pelas setas contínuas na Figura 1) e tendo como retorno o modelo inferido (conforme mostrado pelas linhas tracejadas na Figura 1). A passagem do modelo como parâmetro de entrada é representada pelo seu valor literal, devido à sua simplicidade, garantindo a transparência de localização e minimizando o trabalho de conversão para um protocolo de aplicação.

Serviços de inferências são custosos, para solucionar esta questão, propõe-se a adoção de um *pool* de conexões no motor de inferências, a fim de limitar o seu acesso e consequentemente preservar a disponibilidade deste recurso. Através da definição da quantidade máxima de requisições, será possível que os serviços de inferências operem dentro dos limites que o *hardware* da plataforma permite. Só haverá aumento no número de acessos depois que novos recursos de *hardware* forem adicionados e a quantidade máxima de requisições ao *pool* for aumentada, provendo um estágio inicial em relação à escalabilidade.

3. Experimentos e Resultados

A presente definição do motor de inferências como serviço está sendo aplicada em um sistema de gerência autônoma de redes, e encontra-se implementada em linguagem Java, utilizando as ferramentas descritas acima.

Durante os testes, um modelo didático de ontologia de redes foi desenvolvido na ferramenta Protégé (Stanford Center for Biomedical Informatics Research, 2013). Um *bundle* cliente foi criado para consumir os serviços propostos. A partir do cliente foram realizadas as chamadas aos serviços desenvolvidos sobre a ontologia de teste e os resultados das inferências foram salvos. Para uma comparação, as mesmas inferências foram realizadas na ferramenta Protégé. Finalmente, os resultados do *bundle* cliente e os da ferramenta Protégé foram comparados e demonstraram ser equivalentes, comprovando a consistência das informações encontradas.

A solução trouxe como resultado a diminuição da complexidade em lidar com múltiplas instâncias de *reasoners* e seus artefatos, uma vez que foi abstraído o acesso à

ferramenta de inferências Pellet, permitindo a gestão mais simplificada de seus recursos. Além disso, promoveu-se o reuso através da definição de um componente responsável por serviços de inferência, ao invés da replicação desses serviços em cada aplicação.

4. Considerações Finais

O trabalho apresenta os benefícios da modularidade de motores de inferência em plataforma OSGi para diminuição da complexidade e reuso em sistemas que requerem a manipulação de artefatos semânticos. É importante destacar que a centralização de serviços de inferências acrescenta um ponto de falha para sistemas baseados em ontologias, uma vez que a indisponibilidade do motor de inferências compromete o funcionamento de todos os requisitantes do serviço. Por isso, a arquitetura apresentada deve ser aprimorada visando um melhor desempenho e qualidade dos resultados.

Um trabalho correlacionado é o OWLTools, que efetua inferências do tipo checagem de consistência, descoberta de subclasses ou superclasses, dentre outras. O que difere desta abordagem é o fato de tratar-se de uma API independente destinada a oferecer suporte completo para artefatos OWL, com funcionalidades além dos *reasoners*, enquanto a solução proposta destina-se a oferecer apenas serviços de inferências em um ambiente modular.

Futuros trabalhos podem contemplar a possibilidade de suportar a adoção de outros *reasoners* além do Pellet sob escolha parametrizável. Também há possibilidade de exportar as funcionalidades de um motor de inferências para outras plataformas além de OSGi, através da definição de outros artefatos modulares específicos que implementem a mesma interface de serviço. Por exemplo, utilizar o acesso via WebServices ou RMI (*Remote Method Invocation*). Por fim, é possível aplicar a abordagem proposta neste trabalho para a implementação de outros módulos de igual comportamento, como, por exemplo, um motor de persistência de ontologias.

Referências

- Clark & Parsia (2011). “Pellet Help”, <http://clarkparsia.com/pellet/docs/>, Dezembro.
- The Eclipse Foundation (2013). “Equinox Resources”, <http://www.eclipse.org/equinox/resources.php>, Janeiro.
- Gruber, T. (1993). “A Translation Approach to Portable Ontology Specifications”. Knowledge Specifications, v. 5, n. 1, p. 199-220, Junho.
- Hewitt, E. (2009). Java SOA Cookbook, 1ª edição.
- Lindörfer, F. (2010). “Semantic Web Frameworks”, http://cswwwarchiv.cs.unibas.ch/lehre/hs10/cs341/_Downloads/Workshop/Reports/2010-HSDIS-F_Lindorfer-Semantic_Web_Frameworks-Report.pdf, Dezembro.
- Monteiro, M. E. (2010). “Uma Proposta de Serviços Semânticos Relacionada ao Autogerenciamento em Redes Ópticas de Transporte”. Tese (Doutorado em Engenharia Elétrica – UFES).
- OSGi Alliance (2012). “The OSGi Architecture”, <http://www.osgi.org/Technology/WhatIsOSGi>, Dezembro.
- Stanford Center for Biomedical Informatics Research (2013). “Protégé user documentation”, <http://protege.stanford.edu/doc/users.html>, Dezembro.

Um Sistema para Visualização Tridimensional das Mamas a partir da Curva do Limite Inferior Frontal das Mamas

Geann V. Marozzi, Karin S. Komati

Coordenadoria de Informática – Instituto Federal do Espírito Santo (IFES) Campus Serra - Rodovia ES-010, Km 6.5 – 29.164-231 – Serra – ES – Brasil

geann.vm@gmail.com, kkomati@ifes.edu.br

Abstract. *In this paper we propose a system for visualizing three-dimensional trunk of women by modeling the breasts under their front curve lower bound. The proposed architecture is divided into two sequential stages: the first merges a reference model with a curve, creating a two-dimensional image which is then passed to a tool for generating 3D models from 2D images.*

Resumo. *Neste trabalho propõe-se uma solução para a visualização tridimensional de troncos de mulheres, modelando as mamas de acordo com a sua curva do limite inferior frontal. A arquitetura proposta é dividida em dois estágios sequenciais: na primeira mescla-se um modelo de referência com uma curva, criando-se uma imagem bidimensional que é então passada para uma ferramenta de geração de modelos 3D a partir de imagens 2D.*

1. Introdução

A possibilidade de visualizar e manipular interativamente imagens em 3D permite a compreensão e análise de informações de natureza espacial com mais eficiência, explorando a capacidade humana de raciocinar visualmente. Grandezas como profundidade, largura e altura dos elementos da imagem, podem indicar informações como: seus lugares mais quentes, mais altos ou mais distantes; ou diferenças de tamanho; ou informações absolutas como distância, posição e tamanho dos objetos. Com um fácil controle de navegação para rotacionar, transladar, afastar ou aproximar a imagem, é possível amplificar as informações extraídas, o que muitas vezes não seria possível com uma imagem em 2D [Assis e Vieira, 2005].

Este trabalho propõe uma arquitetura para gerar uma visualização de imagens de 3D de mamas a partir da curva do limite inferior frontal. Essa modelagem tem diversas aplicações importantes como geração de malhas para uso em métodos numéricos de análise e na visualização do resultado de uso de próteses ou reconstrução por cirurgias reparadoras [Castro et. al., 2009].

2. Solução Proposta

O sistema proposto possui dois estágios de processamento, o primeiro é o *Morphing* e o segundo é a geração da modelagem 3D, conforme Figura 1.

Morphing de imagens é uma técnica de processamento de imagem usada para calcular uma transformação, ou "metamorfose", a partir de uma imagem para outra. Um software *morphing* permite que você transforme uma imagem em outra, gerando uma

seqüência de imagens intermediárias de transição. O aplicativo utilizado foi o “Sqirlz Morph” [Xiberpix, 2012], que é *freeware*. O *morphing* recebe como entrada um modelo do tronco corporal (foto bidimensional de um manequim) e a curva do limite inferior frontal de uma pessoa real (imagem bidimensional).

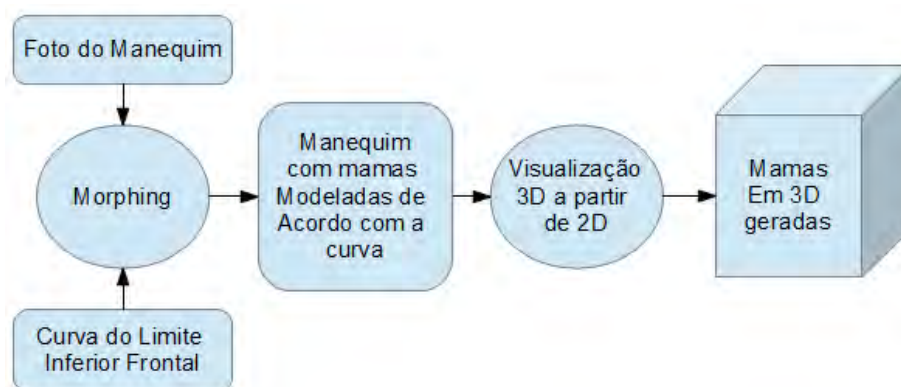


Figura 1. A arquitetura do sistema.

A imagem utilizada do *morphing* é a imagem do meio do processo de transição que representa o manequim já moldado de acordo com a curva. É esta imagem bidimensional que será modelada em 3D. Esta ferramenta foi totalmente desenvolvida para este fim e será descrita mais detalhadamente a seguir.

2.1. Ferramenta de Visualização de Imagens em 3D a partir de 2D

A ferramenta foi desenvolvida em linguagem Java, usando a biblioteca JOGL [JOGL, 2012], que é um projeto independente de código aberto sob licença BSD. JOGL é um *wrapper* (invólucro) que permite que o OpenGL [Kronos Group, 2012] seja usado na linguagem Java.

A concepção da ferramenta é tratar a cor de cada ponto da imagem em 2D como a altura da visualização da imagem em 3D, cada cor da imagem colorida é inicialmente convertida para a escala de cinza. Além disso, a ferramenta disponibiliza a configuração de indexação de cores RGB, e utiliza este índice como altura. Um exemplo é mostrado na Figura 2.

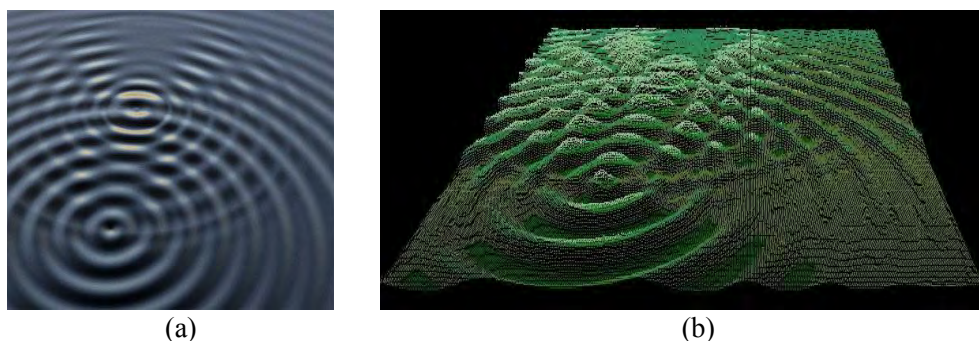


Figura 2. Entrada e saída da Ferramenta de Visualização 3D a partir de imagens 2D, (a) um exemplo de uma imagem de entrada colorida e (b) uma visualização poligonal 3D com alteração da escala de cores associada à imagem de entrada.

O resultado tridimensional pode ser visualizado em quatro formatos de renderização, imagens formadas por pontos, triângulos, polígonos aramados e polígonos com preenchimento. A grande vantagem desta abordagem é a simplicidade da ideia e um resultado visual bem interessante, desde que a imagem de entrada apresente variações de cor que forneçam a percepção de profundidade.

3. Experimentos e Resultados

Para os nossos experimentos, utilizamos como imagens das curvas do limite inferior frontal, as imagens da base de dados de imagens de mamografia térmicas [Visual Lab, 2012]. A base de dados é composta por 285 imagens, a Figura 3(a) mostra um exemplo de uma imagem de mamografia térmica e em (b) a anotação associada da curva feita por um especialista.



Figura 3. (a) Imagem IR_0089.jpg (b) IR_0089-Mask.png, retiradas de [Visual Lab, 2012].

A foto do manequim usada como modelo de referência é mostrada na Figura 4(a). O fundo da foto foi retirado e a sua visualização 3D em malha poligonal é mostrada na Figura 4(b), e sua visualização 3D poligonal com preenchimento é mostrada na Figura 4(c).

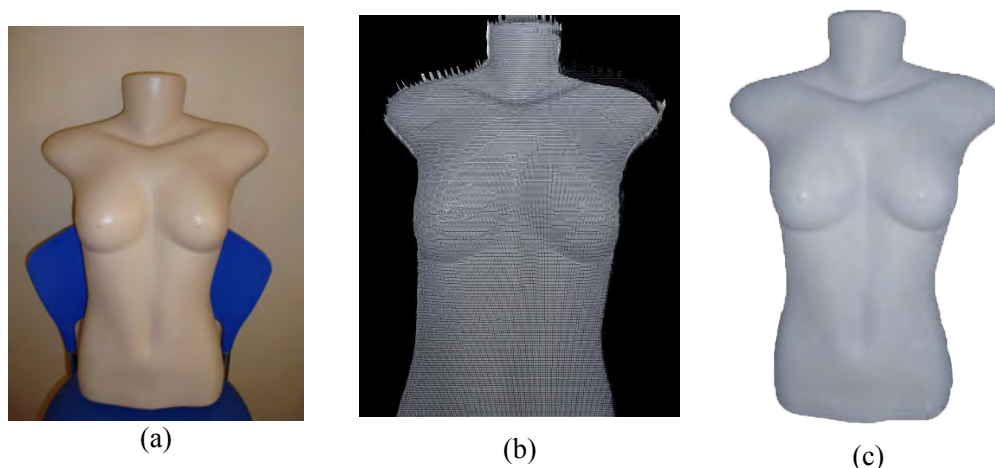


Figura 4. Criação do modelo do tronco corporal, em (a) a foto de um manequim e em (b) o resultado aramado em 3D gerada através de polígonos.

Já com as imagens de entrada: a curva do limite inferior frontal das mamas e de um modelo de referência, o processo de *morphing* é executado e o resultado médio é

modelado em 3D. A Figura 5 mostra um exemplo dos testes realizados, onde (a) é a imagem de entrada de uma curva, (b) é a parte do tronco do modelo de referência e (c) é o resultado em 3D.

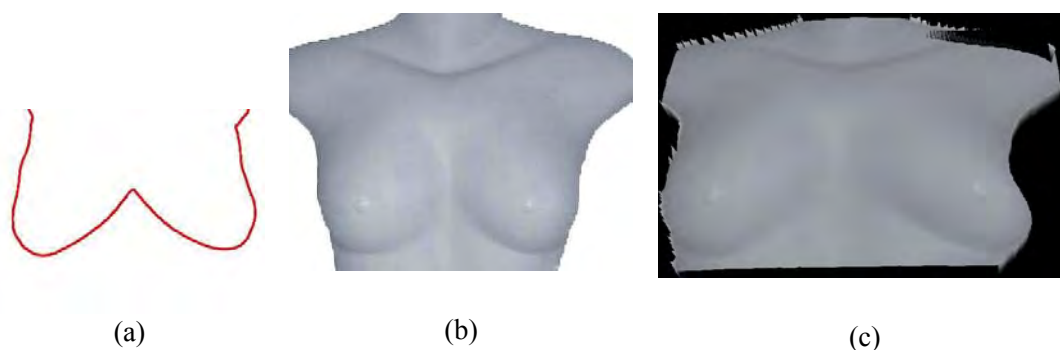


Figura 5. Um exemplo de imagens de entrada (a) curva do limite inferior frontal das mamas (b) manequim e (c) o resultado final da modelagem em 3D.

4. Considerações Finais

Este trabalho está em fase de desenvolvimento, e dispõe-se de resultados parciais. É possível melhorar os resultados para que fiquem mais próximos das imagens reais dos pacientes. Acredita-se que a inclusão de imagens com o perfil lateral das mamas irá melhorar em muito o resultado da modelagem 3D.

Um trabalho futuro é incluir o posicionamento de tumores que forem descobertos nas análises das imagens termográficas de mamas, auxiliando o médico em seu diagnóstico.

Referências

- Assis, S. E., Vieira, R. M. (2005). “View 3D”. Trabalho de Conclusão de Curso em Ciência da Computação. Faculdade Vitoriana de Tecnologia (FAVI). Vitória, ES.
- Castro, F. J., Vasconcelos, S.; Serrano, R. C., Motta, L. S., Menezes, P. M.; Bezerra, L. A., Lima, R. C. F. e Conci, A. (2009) “Um Sistema para pré-processamento de Imagens Térmicas e Modelagem Tridimensional Aplicadas à Mastologia”. In: Anais do II Encontro Nacional de Engenharia Biomecânica: ENEBI 2009, 2009, Florianópolis, v. 1, p. 83-84.
- JOGL (2012). “Java Binding for the OpenGL API”. <http://jogamp.org/jogl/www/>, Agosto.
- Kronos Group (2012). “OpenGL: The Industry's Foundation for High Performance Graphics”. <http://www.opengl.org/>, Agosto.
- Visual Lab (2012). “Processamento e Análise de Imagens Aplicadas à Mastologia”. <http://visual.ic.uff.br/proeng/>, Agosto.
- Xiberpix (2012) "Sqirlz Morph". <http://www.xiberpix.net/SqirlzMorph.html>, Agosto.

Detecção, caracterização e recuperação de faces

Cleriston Aparecido Dantas¹, Marcos Aurélio Batista¹

¹Departamento de Ciência da Computação
Universidade Federal de Goiás (UFG)
Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário
75704-020 – Catalão – GO – Brasil

cleriston.tato@gmail.com, marcos@catalao.ufg.br

Abstract. *In the last years, the use of less invasive technology has been widely used both in biometric access control as security and document validation. One of the primary methods is the facial recognition, it does not require the intervention of the individual. The process of face recognition in images or videos is provided by the steps of locating faces, classify them, and finally identifying individuals through a database of faces. This work consisted of a bank faces and a face recognition system with skin segmentation, location, classification and comparison of faces. Some tests were performed in order to test the efficiency of the developed system, fifteen tests performed on segmentation of skin and six tests in the identification of individuals.*

Resumo. *Nos últimos anos o uso de tecnologia menos invasivas tem sido amplamente utilizados, tanto no controle de acesso biométrico quanto na segurança e validação de documentos. Um dos principais métodos é o de reconhecimento facial, pois não necessita da intervenção do indivíduo. O processo de reconhecimento facial em imagens ou vídeos é dado pelas etapas de localizar as faces, classificá-las e por fim identificar indivíduos, por meio de um banco de dados de faces. Neste trabalho, constituiu-se um banco de faces e um sistema de reconhecimento facial, com segmentação de pele, localização, classificação e comparação de faces. Alguns testes foram realizados com o intuito de testar a eficiência do sistema desenvolvido, realizou-se quinze testes na segmentação de pele e seis testes na identificação dos indivíduos.*

1. Introdução

Tem-se usado diversos recursos tecnológicos para a identificação de indivíduos. O uso de impressões digitais e DNA ainda são os principais métodos de identificação, porém, muitas vezes intrusivos, necessitando da colaboração do usuário. Por isso o uso de tecnologias não invasivas, como o reconhecimento de faces, tem se tornado cada vez mais comuns.

Os primeiros estudos de reconhecimento facial datam na década de 1950, empregados em estudos de psicologia [Bruner and Tagiuri 1954], na década de 1970 iniciou-se o uso de máquinas automáticas para reconhecimento facial, como se vê nos trabalhos de [Kanade 1977] e [Kelly 1970].

Nestes últimos 40 anos as pesquisas sobre o reconhecimento facial sofreram grande avanço, principalmente ao se fazer uso de abordagens holísticas e clássicas como

eigenfaces [Turk and Pentland 1991a], *fisherfaces* [Etemad and Chellappa 1997] e mais recente o uso de modelos tridimensionais [Campos 2001] [Colombo et al. 2005].

O reconhecimento automatizado pode ser útil em diversas aplicações, como: controle de acesso, rastreamento de suspeitos, realidade virtual, passaportes, interação homem-computador, em aplicações sociais e até mesmo usado como uma ferramenta de auxílio aos estudos de psicologia [Zhao et al. 1998].

O reconhecimento facial se dá por três etapas: detecção da face na imagem ou vídeo, caracterização desta face encontrada e por fim a recuperação da mesma [Wu et al. 2012]. No desenvolvimento deste trabalho será discutido cada etapa.

2. Metodologia

Como dito, o trabalho é dividido basicamente em três passos principais, dado uma imagem de entrada, detectar a face, extrair características da mesma, essas características são salvas em um arquivo que servirá como uma base de dados para que no passo seguinte, ao entrar com uma imagem nova, identificar o indivíduo presente nela. Todas as rotinas foram implementadas em C++ e usando a biblioteca OpenCV. A Figura 1 ilustra o processo.

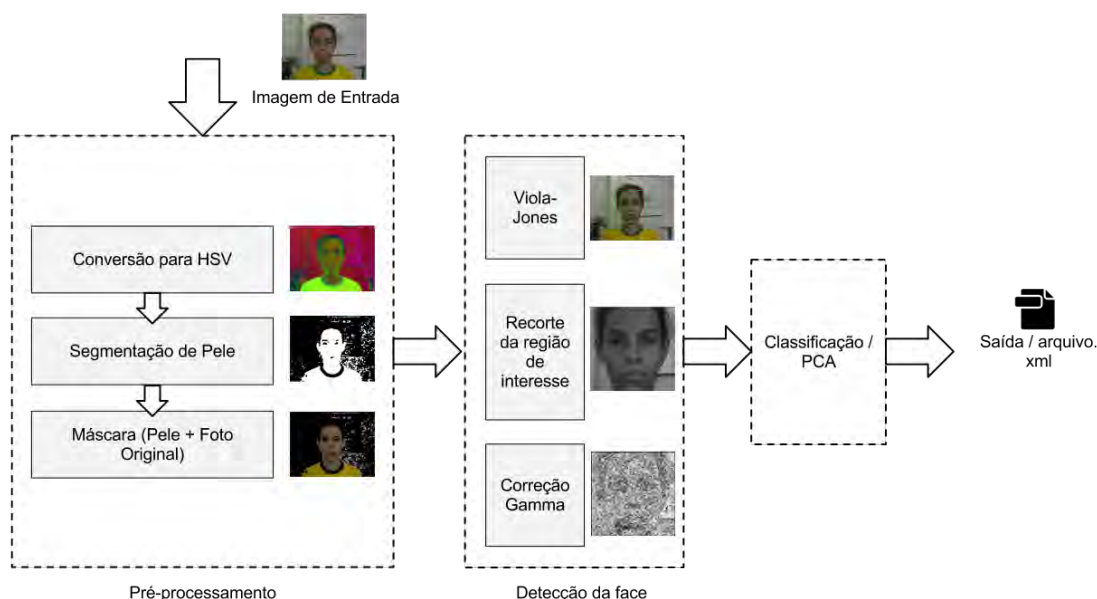


Figura 1. Etapas adotadas para o reconhecimento facial neste trabalho.

2.1. Pré Processamento, segmentação da pele

Antes de iniciar o processo de detecção, fez-se uso de uma etapa de pré-processamento das imagens, no qual a partir de uma imagem de entrada, esta, foi convertida de RGB (*Red*, *Blue* e *Green*) para HSV (*Hue*, *Saturation* e *Value*), esta medida é necessária para a segmentação da pele, pois ela é melhor representada pelo canal *Hue*. Após quinze testes com 251 imagens presentes no banco criado, o melhor limiar encontrado foi entre 3 e 55, o processo envolveu a variação dos limiares, escolhemos aquele que obteve a melhor segmentação de pele sem interferir na detecção de face e apresentar o mínimo de falsos

positivos, como na Figura 2. Após realizada a distinção de pele e não pele, gerou-se uma imagem contendo a imagem original e a segmentação.

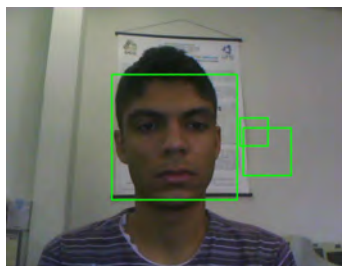


Figura 2. Exemplo de falsos positivos, foi detectado faces onde não existe.

2.2. Detecção da face

Após isso, a detecção da face é realizada na imagem gerada, nesta etapa destacam-se métodos por atributos de gabor, métodos tridimensionais, holísticos, neste trabalho o uso de características de *haar*, por meio do algoritmo de Vioja-Jones [Viola et al. 2003] foi feito. Como o banco de faces não foi obtido em um ambiente totalmente controlado, capturou-se sem intenção imagens com os indivíduos em distâncias diferentes, portanto todas as faces passaram por um reajuste de tamanho no rosto encontrado, limitando um quadrado de 150px, o reajuste evita que na etapa próxima haja alguma incompatibilidade por tamanho das imagens.

2.3. Caracterização das faces

Na etapa seguinte, são classificadas todas as faces encontradas, diante de alguns métodos na literatura como, *Fisherface* [Jing et al. 2006], *Independent Component Analysis* (ICA) [Bartlett et al. 2002], *Linear discriminant analysis* (LDA) [Lu et al. 2003], *Evolutionary Pursuit* (EP) [Zhao et al. 2006], *Elastic Bunch Graph Matching* (EBGM) [Bolme 2003] e o modelo tridimensional de Li e Zeng [Li et al. 2012], optou-se por implementar no programa o *Eigenface*, conhecido também como *Principal Components Analysis* (PCA) [Turk and Pentland 1991b] [Sevcenco and Lu 2010], pelo uso do método original inventado por Karl Pearson [Pearson 1901], é um dos métodos mais clássicos e que resulta em ótimos resultados.

A ideia do PCA é bem simples, ele reduz as variáveis originais de uma imagem em um conjunto mínimo, mas que contenham o máximo de informação possível da imagem original, assim obtém-se uma matriz reduzida com todas as características principais da matriz completa.

Após a extração dessas características, foi criada uma imagem média e todos os dados resultantes foram armazenados em uma estrutura de um arquivo XML, juntamente com o número da foto e indivíduo. Este arquivo fonte contém os dados de todas as pessoas do banco de dados constituído.

2.4. Recuperação da face

A etapa final desse sistema consiste em entrar com uma nova imagem que pasará por todas as etapas anteriores e gerada sua estrutura individual, os dados desta

estrutura serão comparadas com o arquivo fonte. Esta comparação foi realizada por dois métodos de cálculo de distância. O primeiro teste com a distância Euclidiana [Gonzalez and Woods 2000], e o segundo com a distância de Mahalanobis [Chan et al. 2012]. Para cada ponto da imagem de consulta é comparada com pontos de cada imagem, o conjunto de pontos mais próximos determina se o indivíduo é ou não semelhante.

A distância Euclidiana entre P_1 e P_2 é dado pela equação 1 [Gonzalez and Woods 2000].

$$D_e(p_1.p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

E a distância Mahalanobis, é dada pela equação 2, onde μ_x é a imagem média [Chan et al. 2012].

$$D = \sqrt{(x - \mu_x)^t \sum_x^{-1} (x - \mu_x)} \quad (2)$$

O sistema retorna o indivíduo mais parecido com o indivíduo consulta, mesmo que este não se encontra no banco de faces, uma informação importante é o fato da imagem da face de entrada poder ser em três diferentes posições, além de uso ou não de óculos, isto foi possível pelo modo que o banco foi constituído.

2.5. Criação do banco de faces

O banco de faces usado neste trabalho foi constituído com base em 7 imagens em diferentes poses e uso ou não de óculos, conforme a Figura 3. As imagens foram capturadas por meio de uma *webcam* comum, presente na maioria dos *notebooks*, a resolução adotada foi 640 x 480 px. Nenhum critério foi considerado para a constituição deste banco de faces, todas as pessoas foram voluntárias, assinando inclusive um termo de autorização de uso de suas imagens pelo sistema capturado.

O sistema em *flash* capturava as imagens nas diferentes poses e às armazenava em um banco MySQL, um sistema para gerenciar estas imagens também foi desenvolvido Figura 4, para que pudesse realizar buscas operações nos registros.

3. Resultados

O uso da segmentação de pele tem objetivo diminuir a taxa de falsos positivos, no qual realmente o fez, reduzindo de 32 falsos positivos na detecção de faces sem uso de segmentação de pele, para 8 falsos positivos, o resultado foi obtido pelo limiar entre 3 e 55, todos os testes para cada limiar pode ser observado na Tabela 1, em alguns casos como no limiar entre 3 e 25, as taxas de faces não detectadas foram muito altas.

Para a identificação dos indivíduos, realizamos seis testes, os testes consistiam em treinar 29 imagens, uma de cada indivíduo, e testar outras 58 não presentes no treinamento, foram três conjuntos distintos. A comparação dos indivíduos foram realizadas com uso da distância de Euclides e a de Mahalanobis, ver Tabela 2.

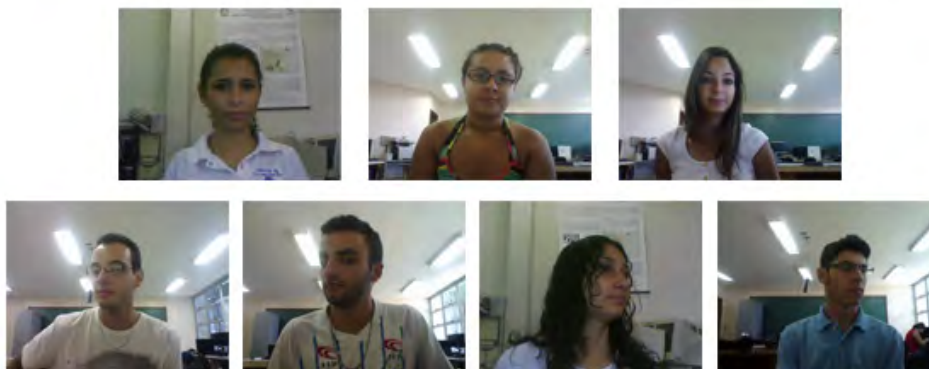


Figura 3. Diferentes posições em que o banco de dados foi constituído.

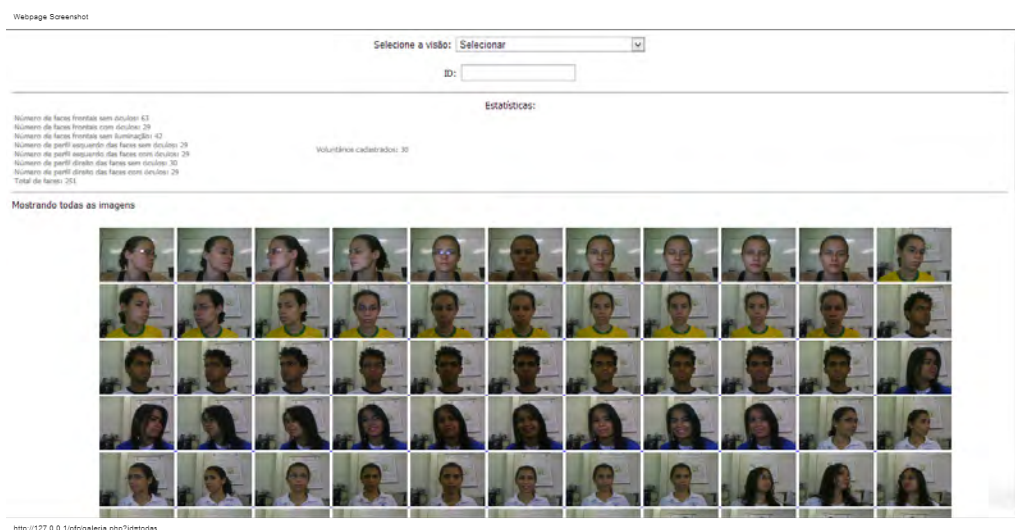


Figura 4. Sistema de gerenciamento do Banco de Faces.

Podemos observar que os resultados com uso da distância de Mahalanobis é levemente superior ao método de Euclides.

4. Considerações Finais

O sistema aqui desenvolvido propunha o reconhecimento de indivíduos por meio do uso de alguns métodos e abordagens clássicas, usou-se segmentação de pele, método de Viola-Jones para detecção de faces, o PCA para extração de características e os métodos de Mahalanobis e de Euclides para cálculo de distância.

Pode se observar que o sistema desenvolvido obtém um bom resultado para a segmentação de pele, sem uma grande interferência na detecção do rosto, além de diminuir a taxa de falsos positivos.

Na identificação dos indivíduos o modelo de Mahalanobis se mostrou superior ao modelo Euclidiano, obtendo taxas de até 77,59% de acerto, o que pode ser considerado

Limiar	Faces detectadas	Certos	Erros (falsos positivos)	Não detectadas
sem limiar	223	191	32	60
0-30	229	209	20	42
0-40	236	208	18	43
0-50	206	201	5	50
0-60	212	198	14	53
0-70	231	196	35	55
3-25	120	108	12	143
3-35	204	198	6	53
3-45	204	198	6	53
3-55	223	215	8	36
3-65	226	197	29	54
10-40	215	199	16	52
10-50	207	194	13	57
10-60	221	193	28	58
10-70	216	180	36	71

Tabela 1. Comparação entre os diferentes limiares testados

Conjunto	Euclidiana	Mahalanobis
Treinamento de faces com óculos	74,14%	77,59%
Treinamento de faces com iluminação	53,45%	63,80%
Treinamento sem iluminação	58,62%	67,24%

Tabela 2. Resultados da identificação dos indivíduos, com uso de dois modelos de cálculo de distância

razoável, tendo em mente que o banco foi constituído em ambiente não controlado.

A principal dificuldade foi em constituir o banco de dados, pois muitas pessoas ficam com receio de ser voluntário em um trabalho, ainda mais quando sua foto está envolvido.

Para os próximos trabalhos, a implementação de outros métodos de extração de características serão realizados, o banco de faces será aumentado e outros métodos de comparação serão testados.

Referências

- Bartlett, M., Movellan, J., and Sejnowski, T. (2002). Face recognition by independent component analysis. *Neural Networks, IEEE Transactions on*, 13(6):1450–1464.
- Bolme, D. (2003). *Elastic bunch graph matching*. PhD thesis, Colorado State University.
- Bruner, I. and Tagiuri, R. (1954). The perception of people. *Handbook of Social Psychology*, 2:634–654.
- Campos, T. E. (2001). Técnicas de seleção de características com aplicações em reconhecimento de faces. Master's thesis, Instituto de Matemática e Estatística, University of São Paulo.

- Chan, C., Tahir, M., Kittler, J., and Pietikainen, M. (2012). Multiscale local phase quantisation for robust component-based face recognition using kernel fusion of multiple descriptors.
- Colombo, A., Cusano, C., and Schettini, R. (2005). Tri-dimensional face detection and localization. In *Proceedings of SPIE - The International Society for Optical Engineering*, volume 5670, pages 68–75.
- Etemad, K. and Chellappa, R. (1997). Discriminant analysis for recognition of human face images. *Journal of the Optical Society of America A*, 14(8):1724–1733.
- Gonzalez, R. and Woods, R. (2000). *Processamento de imagens digitais*. Edgard Blucher.
- Jing, X., Wong, H., and Zhang, D. (2006). Face recognition based on 2d fisherface approach. *Pattern Recognition*, 39(4):707–710.
- Kanade, T. (1977). *Computer recognition of human faces*, volume 47. Birkhäuser.
- Kelly, M. (1970). Visual identification of people by computer. Technical report, DTIC Document.
- Li, C., Mu, Z. ., and Zeng, H. (2012). Human identification based on 3d ear modeling using a single image. *International Journal of Advancements in Computing Technology*, 4(23):88–96.
- Lu, J., Plataniotis, K., and Venetsanopoulos, A. (2003). Face recognition using kernel direct discriminant analysis algorithms. *Neural Networks, IEEE Transactions on*, 14(1):117–126.
- Pearson, K. (1901). *Principal Components Analysis*, page 559. Cited By (since 1996): 1.
- Sevcenco, A.-M. and Lu, W.-S. (2010). Perfect histogram matching pca for face recognition. *Multidimensional Systems and Signal Processing*, 21:213–229.
- Turk, M. and Pentland, A. (1991a). Face recognition using eigenfaces. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 586–591.
- Turk, M. and Pentland, A. (1991b). Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR'91., IEEE Computer Society Conference on*, pages 586–591. IEEE.
- Viola, M., Jones, M., and Viola, P. (2003). Fast multi-view face detection. In *Proc. of Computer Vision and Pattern Recognition*. Citeseer.
- Wu, Y.-M., Wang, H.-W., Lu, Y.-L., Yen, S., and Hsiao, Y.-T. (2012). Facial feature extraction and applications: A review. *Intelligent Information and Database Systems*, pages 228–238.
- Zhao, Q., Lu, H., and Zhang, D. (2006). A fast evolutionary pursuit algorithm based on linearly combining vectors. *Pattern Recognition*, 39(2):310–312.
- Zhao, W., Chellappa, R., and Krishnaswamy, A. (1998). Discriminant analysis of principal components for face recognition. In *Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference on*, pages 336–341. IEEE.

Estudo de Caso Usando o *Framework* Robot Operating System (ROS)

Iohan G. Vargas¹, Dalton Matsuo Tavares¹

¹Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão – GO – Brasil
iohanufg@gmail.com, dalton.tavares@catalao.ufg.br

Abstract. *This paper briefly describes the ROS framework and its functionalities, presenting its strengths and demonstrating its usability in a practical application, whose goal is to read the touch sensor of the LEGO® Mindstorms® Education NXT 9797 kit. By means of a rich toolset offered by the ROS framework, it is possible to efficiently and promisingly circumvent the limitations of the native LEGO software, using a computer as an extension, which will perform additional processing and therefore, allow the creation of more complex prototypes.*

Resumo. *Este artigo descreve de forma sucinta o framework ROS e suas funcionalidades, apresentando pontos positivos e demonstrando sua usabilidade em uma aplicação prática cujo objetivo é a leitura do sensor de toque do KIT LEGO® Mindstorms® Education NXT 9797. Por meio de um rico conjunto de ferramentas oferecido pelo framework ROS, é possível contornar de forma eficiente e promissora as limitações do software nativo do KIT LEGO®, utilizando como extensão de processamento um computador, o qual deve realizar processamento adicional e conseqüentemente, permitir a criação de protótipos mais complexos.*

1. Introdução

Robot Operating System (ROS) é uma plataforma de software de código aberto projetado para suportar uma nova geração de robôs pessoais. A comunidade ROS desenvolveu o *framework* para uma ampla gama de plataformas, disponibilizando-o para diferentes sistemas operacionais (SOs), como Windows, Linux e Mac OS. O *framework* ROS é um rico conjunto de ferramentas que gerencia de forma eficiente a mediação entre Sistema Operacional (SO) e demais aplicações, fornecendo bibliotecas reutilizáveis e ferramentas que são projetadas para trabalhar de forma independente, com o intuito de ajudar desenvolvedores a criarem aplicações no campo da robótica. Pode-se afirmar que o ROS representa uma coleção muito útil de softwares voltados exatamente para ajudar na percepção, controle e modelagem dos dispositivos de hardware que serão lidos, fornecendo serviços que são esperados de um SO, incluindo a abstração do hardware em baixo nível, o que permite a leitura e o controle do mesmo, passagem de mensagens entre processos e gerenciamento de pacotes [Quigley et al. 2009].

O ROS possui uma ampla documentação, em sua maioria em língua inglesa¹. Na página da comunidade ROS, existe um extenso material de apoio, oferecendo suporte a

1 Disponível em <http://www.ros.org/wiki/> Acesso em: 17/01/2013.

compartilhamento de códigos fonte e experiências realizadas com o *framework* ROS, além de artigos científicos que envolvem sua utilização em diversas áreas do conhecimento. A comunidade ROS possibilita ao usuário fazer parte da mesma, mediante um cadastro pessoal no site. Dessa forma, o usuário pode deixar eventuais perguntas e/ou dúvidas que poderão ajudá-lo a solucionar algum problema detectado².

Este software é uma ferramenta de código aberto (*open source*), com todo seu código do núcleo licenciado pela licença BSD (*Berkeley Software Distribution License*). Paralelamente o sistema também é projetado para ser capaz de incluir componentes escritos sob várias licenças públicas da GNU (*GNU general public license* ou GPL³) [Cousins 2010b]. Possui uma estrutura distribuída de processos, a qual apoia a reutilização de código em robótica. Suas bibliotecas possuem funcionalidade para diversas linguagens de programação que serão mencionadas na seção 2. Este *framework*, originado da *Stanford University e Research Institute*, é considerado apropriado para grandes processos de desenvolvimento e é atualmente apoiado pela *Willow Garage*.

No contexto desse artigo, o software ROS será usado com o kit LEGO® Mindstorms® Education NXT 9797, com o intuito de compor um estudo de caso visando a leitura de sensores. Para oferecer melhores esclarecimentos para a compreensão do estudo de caso proposto, a seção 2 irá apresentar uma breve descrição do *framework* ROS e de suas funcionalidades, a seção 3 irá apresentar o estudo de caso pretendido e a seção 4 as conclusões e propostas de trabalhos futuros.

2. Robot Operating System (ROS)

O *framework* ROS foi projetado para atender a um conjunto específico de desafios encontrados durante o desenvolvimento em grande escala de robôs de serviço. O ROS é muito mais do que apenas um serviço oferecido a robôs móveis e de manipulação [Alexander et al. 2012].

Uma aplicação ROS é um grafo de nodos, os quais representam um processo no *framework* ROS. Os nodos podem residir em uma mesma máquina ou em máquinas diferentes. A comunicação entre nodos é feita usando dois protocolos, os tópicos (*topics*), protocolo assíncrono que implementa um fluxo de dados e troca de mensagens, possibilitando que múltiplos *publishers/subscribes* concorram a um único tópico, onde a ordem das mensagens é irrelevante e os serviços (*services*), protocolo síncrono que é definido como uma chamada de procedimento remoto, a qual espera por uma mensagem de resposta, sendo possível ter vários clientes simultâneos. Ambos utilizam como diretório mestre, o ROS *Master*⁴.

Sua organização como *framework* pode ser resumida em um serviço ponto-a-ponto (*peer-to-peer*), baseado em ferramentas (*tools-based*), aplicável a múltiplas linguagens (*multi-lingual*), leve (*thin*), livre e de código aberto. Contudo, segundo Quigley et al. (2009), nenhum outro *framework* existente possui esse conjunto de critérios de projeto reunidos em uma só ferramenta. Esses critérios podem ser explicados de forma objetiva, como segue [Quigley et al. 2009]:

1. Ponto-a-ponto: um sistema construído usando ROS consiste em uma série de

2 Disponível em <http://answers.ros.org/questions/> Acesso em: 17/01/2013.

3 Disponível em <http://www.nacaolive.com.br/open-source/licenca-gpl/> Acesso em: 20/01/2013.

4 Disponível em <http://barraq.github.com/fOSSa2012/slides.html#slide-13> Acesso em: 19/01/2013.

processos, potencialmente distribuídos em um número diferente de hospedeiros, ligados em tempo de execução do sistema em uma topologia ponto-a-ponto.

2. Multi linguagem: ao escrever algum código, muitas pessoas têm preferências por uma determinada linguagem de programação. Estas podem opções técnicas ou pessoais. Por essa razão, o ROS foi projetado com um idioma neutro, suportando atualmente C, C ++, Python, LISP. Para fornecer o suporte a diversas linguagens de programação, o ROS utiliza uma IDL (*Interface Definition Language*) , que é uma linguagem neutra de definição de interface para descrever as mensagens enviadas entre os módulos.
3. Baseado em ferramenta: com o intuito de gerenciar a complexidade do ROS, existe um projeto de *microkernel*, no qual um grande número de pequenas ferramentas são usadas para criar e executar vários componentes, ao invés de construir um sistema monolítico.
4. Leve: Tendo em vista a reutilização de código, os desenvolvedores do ROS foram encorajados a se fundamentarem em bibliotecas independentes. O seu sistema de compilação executa construções modulares dentro da árvore de códigos-fonte, usando o CMake⁵, o que torna trivial seguir essa ideologia “leve”.
5. *Free and Open-Source*: como mencionado na seção 1, o código fonte completo do ROS está disponível publicamente, permitindo o desenvolvimento de projetos comerciais e não comercial.

3. Estudo de Caso: Leitura de Sensor usando o ROS

Para o desenvolvimento do presente estudo de caso, foi usado o kit LEGO® Mindstorms® Education NXT 9797. É importante registrar que a última versão do ROS (*fuerte*) não oferece suporte para o pacote LEGO® Mindstorms® NXT. Assim, as versões que podem ser usadas com esse kit de desenvolvimento são versões anteriores, como o ROS Electric ou o ROS Diamondback. A versão usada foi o ROS Electric, por ser uma versão estável e adequada a este estudo de caso. Com as ferramentas disponíveis, abre-se um amplo leque de possibilidades para o controle do kit de desenvolvimento. O foco do presente artigo é a demonstração da leitura de sensores do kit LEGO®, em especial o sensor de toque.

3.1. O kit LEGO® Mindstorms® NXT

O kit LEGO® Mindstorms® NXT 2.0 é um kit robótico modular produzido pela LEGO, é constituído por um conjunto de peças de encaixe, acrescido de sensores e motores que são controlados por um processador programável ou bloco inteligente, o qual permite a leitura de sensores e acionamento de motores [Quirino e Gonçalves 2010]. Esse kit é interessante no contexto experimental por ser voltado à educação tecnológica⁶.

O bloco inteligente permite a conectividade via módulo *Bluetooth* ou porta USB 2.0. Disponibiliza três portas de saída e quatro portas de entrada digitais, possui um *display* que mostra as funcionalidades do bloco e um alto-falante. Além do bloco inteligente, o kit LEGO® Mindstorms® Education NXT 9797 possui bateria recarregável de lítio; três servo-motores; dois sensores de toque, um sensor ultra-sônico, um sensor de

5 Cmake é um sistema de automação de compilação de sistemas. Disponível em <http://leandro.setefaces.org/2011/tutorial-cmake-p-i/> Acesso em: 20/01/2013.

6 Disponível em <http://mindstorms.lego.com/en-us/whatisnxt/default.aspx> Acesso em: 17/01/2013.

cor e de intensidade luminosa [Silva e Porto 2011]. Seu hardware apresenta alta durabilidade e facilidade de alteração considerando as peças de encaixes, sendo caracterizado por um custo relativamente baixo em comparação a outros kits (ex. Aldebaran NAO⁷, bioloid⁸ etc).

A interação com o kit LEGO[®] Mindstorms[®] NXT é realizada por meio de uma conexão entre o SO Ubuntu 11.10 e o bloco inteligente, via *Bluetooth* ou cabo USB. Em geral, a interface usual para o ROS é um terminal no qual são executados programas em linhas de comando.

Muito embora o ROS não seja considerado “amigável ao usuário”, as limitações do ambiente de desenvolvimento nativo fornecido com o kit LEGO[®] motivam o seu uso. Dentre as limitações observadas, pode-se citar uma linguagem de programação limitada, a qual não permite, por exemplo, a criação de um servidor baseado em *sockets* TCP no bloco inteligente e não dá suporte a tele-operação de robôs. Também existem limitações considerando o hardware nativo como a baixa capacidade de processamento (Atmel 32 bits ARM7) e pouca quantidade de memória (256 Kbytes FLASH, 64 Kbytes de RAM). Dessa forma, justifica-se o uso do *framework* ROS como forma de superar as limitações observadas, sendo possível criar protótipos mais complexos, mediante o uso de um computador como extensão para realizar processamento adicional.

3.2. Preparação do Espaço de Trabalho

Para instalação do ROS Electric, o site oficial do ROS oferece o procedimento completo para o SO Linux Ubuntu⁹. A versão usada para a confecção deste artigo foi a 11.10 em plataforma de 32 bits. O suporte estável para o ROS é fornecido para Linux Ubuntu, embora existam outros SOs com suporte experimental. Há quatro configurações padrão para a instalação do ROS. A mais apropriada para o estudo de caso pretendido, é a instalação da versão *Desktop-Full Install*, que é a recomendada pelo site do ROS. Para essa instalação é necessário estar conectado a Internet.

Após a instalação bem sucedida do ROS Electric é necessário modificar as configurações do interpretador de comandos usado, nesse caso, o *bash*, de modo que este adicione o ambiente de desenvolvimento do ROS a cada sessão iniciada. Além dos pacotes básicos do ROS Electric, recomenda-se a instalação de algumas ferramentas independentes, as quais poderão ser usadas posteriormente no sistema ROS. São elas o *rosinstall*, usada para adquirir facilmente pacotes e o *roscdep* que permite instalar de forma prática dependências à instalação original do ROS. Em seguida, é necessário criar uma *workspace* pessoal, onde os pacotes desenvolvidos serão criados. Esse procedimento é realizado mediante a execução do comando `rosws init ~/ros_workspace /opt/ros/electric`.

O espaço de trabalho padrão do ROS é configurado em `/opt/ros/electric/setup.bash`. Dessa forma, é preciso mudar esse diretório para a pasta criada anteriormente. Para tanto, basta editar o arquivo `.bashrc` e acrescentar o caminho real da nova pasta de trabalho (ex. `source`

7 Disponível em <http://www.aldebaran-robotics.com/en/> Acesso em: 28/01/2013

8 Disponível em <http://www.trossenrobotics.com/bioloid-and-humanoid-robot-kits.aspx> Acesso em: 28/01/2013.

9 Disponível em: <http://www.ros.org/wiki/electric/Installation/Ubuntu> Acesso em: 17/01/2013.

/home/user/ros_workspace/setup.bash). Além disso, é necessário atualizar a referência do arquivo `setup.bash` (ex. `source ~/ros_workspace/setup.bash`). Em seguida é criada uma pasta específica para uso na criação de pacotes ROS. Essa pasta é colocada dentro do diretório `ros_workspace` (ex. `mkdir ~/ros_workspace/sandbox`). O caminho do pacote deve ser inserido no *path* do ROS (`ROS_PACKAGE_PATH`), com o comando `source /home/user/ros_workspace/setup.sh`. Dessa forma, tem-se o sistema ROS instalado e a pasta pessoal de trabalho criada.

3.3. Adição dos Pacotes Necessários para a operação do kit LEGO®

Os procedimentos de instalação dos pacotes necessários para a operação do kit LEGO® Mindstorms® NXT estão disponíveis no site do ROS¹⁰. O software ROS para LEGO é organizado em pacotes e pilhas, os pacotes contém funcionalidades referentes ao processos do ROS, as pilhas coletam conjuntos de pacotes e juntos oferecem funcionalidade para a comunicação entre nodos, contendo a infraestrutura básica do ROS [Cousins 2010a]. Antes de instalar os pacotes NXT, é necessário fazer algumas configurações.

- A criação de um grupo `lego` e adição do usuário usado ao grupo;
- Criar um arquivo de regras `udev` para o grupo `lego` que acabou de ser criado. Isso pode ser feito de forma simples mediante o comando apresentado na Listagem 1, o qual define regras para o arquivo `lego.rules`. Este arquivo será movido da pasta `/tmp` e será identificado em `/etc/udev/rules.d`, para permitir o acesso via módulo USB ao bloco inteligente. Posteriormente deve-se reiniciar o `udev`.
- Inserir novas fontes para o gerenciador de pacotes, de modo que este possa instalar pacotes do ROS (Listagem 2).
- Instalação do pacote `ros-electric-nxtall` usando o gerenciador de pacotes do Ubuntu.

```
echo "BUS==\"usb\", ATTRS{idVendor}==\"0694\", GROUP=\"lego\",
MODE=\"\" > /tmp/70-lego.rules && sudo mv /tmp/70-lego.rules
/etc/udev/rules.d/70-lego.rules"
```

Listagem 1: Regra criada para o gerenciador de dispositivos udev.

```
Sudo sh -c 'echo "deb http://packages.ros.org/ros/ubuntu
oneiric main" > /etc/apt/sources.list.d/ros-latest.list'
```

Listagem 2: Inserção de nova fonte de pacotes no gerenciador de pacotes do Ubuntu.

Assim, a infraestrutura necessária para a operação do kit LEGO® por intermédio do *framework* ROS está pronta. A seção 3.4 apresentará a leitura de sensores via ROS.

3.4. Leitura de Sensores do LEGO® Mindstorms® NXT

Para leitura de sensores usando o ROS, é necessário certificar-se que a comunicação com

¹⁰ Disponível em: <http://www.ros.org/wiki/Robots/NXT/electric> Acesso em: 17/01/2013.

o bloco inteligente está sendo executada de forma correta. Isso pode ser realizado por meio do comando `roscore` (sem parâmetros), como segue na Figura 1.

```

iohan@iohan-HP-Compaq-8200-Elite-SFF-PC:~$ roscore
... logging to /home/iohan/.ros/log/44aa3a86-280d-11e2-b468-e839351653cf/roslauch
ch-iohan-HP-Compaq-8200-Elite-SFF-PC-2530.log
Checking log directory for disk usage. This may take awhile.
Press Ctrl-C to interrupt
Done checking log file disk usage. Usage is <1GB.

started roslaunch server http://iohan-HP-Compaq-8200-Elite-SFF-PC:51711/
ros_comm version 1.6.6

SUMMARY
=====

PARAMETERS
* /rosversion
* /rostdistro

NODES

auto-starting new master
process[rosmaster]: started with pid [2545]
ROS_MASTER_URI=http://iohan-HP-Compaq-8200-Elite-SFF-PC:11311/

setting /run_id to 44aa3a86-280d-11e2-b468-e839351653cf
process[rosout-1]: started with pid [2558]
started core service [/rosout]

```

Figura 1: Resultado do comando roscore.

A saída do comando apresentada na Figura 1 demonstra como deverá estar o sistema, caso o ROS esteja em execução e a conexão tenha sido realizada com sucesso. A conexão com o bloco inteligente fica evidente considerando a última linha mostrada na Figura 1, (“started core service [/rosout]”). Isso indica que o serviço foi iniciado com sucesso e que a conexão foi estabelecida com o bloco inteligente.

O bloco inteligente possui quatro portas para fixação dos sensores (porta 1, 2, 3, e 4, em sua parte inferior (Figura 2). Na parte superior existem três portas para fixação dos motores e uma porta para conexão via módulo USB. Esta última permite estabelecer conexão com o bloco inteligente, o qual será controlado pelo computador via porta USB. O sensor de toque pode ser conectado à porta 1, como mostrado na Figura 2.

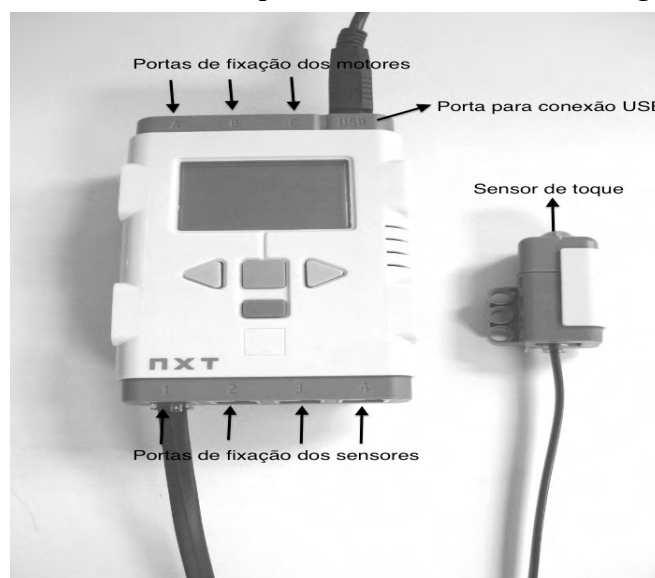


Figura 2: Bloco inteligente e sensor de toque.

A descrição do sensor a ser usado é fornecida ao ROS mediante a especificação de um arquivo do tipo `.yaml`, o qual fornece a descrição do sensor, seu tipo e um identificador. O arquivo `robot.yaml` é especificado para um sensor de toque (Listagem 3).

```
nxt_robot:           // o nome do robô
  - type: touch       // tipo do sensor
    frame_id: touch_frame // identificador para o sensor
    name: my_touch_sensor // nome do sensor
    port: PORT_1        // porta do bloco inteligente
    desired_frequency: 20.0 // frequência definida
```

Listagem 3: Arquivo `robot.yaml` contendo a especificação do sensor de toque.

O próximo passo é a criação de um arquivo de carga com extensão `.launch` para o ROS (ROS *launch*), o qual irá iniciar o pacote NXT-ROS (Listagem 4).

```
<launch>
  <node pkg="nxt_ros" type="nxt_ros.py" name="nxt_ros"
    output="screen" respawn="true">
    <rosparam command="load" file="$(find learning_nxt)/
      robot.yaml" /> </node>
</launch>
```

Listagem 4: Arquivo `robot.launch`.

Agora se faz necessário testar as configurações feitas anteriormente. Assim verifica se o bloco inteligente está se comunicando com o sensor pretendido. Para isto, executa-se o comando `roslaunch robot.launch`. O resultado é a inicialização do serviço NXT-ROS e a criação de uma instância do sensor de toque denominada “`my_touch_sensor`” na porta 1.

```
SUMMARY =====
PARAMETERS
  * /nxt_ros/nxt_robot NODES /
    nxt_ros (nxt_ros/nxt_ros.py)
starting new master (master configured for auto start)
process[master]: started with pid [11699]
ROS_MASTER_URI=http://bvo:11311/
process[rosout-1]: started with pid [11712]
started core service [/rosout]
process[nxt_ros-2]: started with pid [11724]
[INFO] 1282349960.808368: Creating touch with name
my_touch_sensor on PORT_1
```

Listagem 5: Resultado do comando `roslaunch` com as configurações pré-definidas.

A visualização dos sensores pode ser realizada mediante o uso do comando `rostopic list`. Para este estudo de caso, se todos os passos descritos foram executados de forma correta o sensor `my_touch_sensor` aparecerá na lista de sensores em execução. Para verificar o acionamento do sensor de toque execute o seguinte comando `roslaunch my_touch_sensor.py`, no terminal aparecerá os resultados, quando o sensor de toque for pressionado será visualizado `TOUCH: TRUE`, caso contrário `TOUCH: FALSE`.

A visualização dos demais sensores do kit LEGO® Mindstorms® NXT segue os mesmos procedimentos considerando a criação de arquivos de configuração `.yaml` e `.launch`. Vale observar que ao se usar vários sensores paralelamente, pode ser criado apenas um arquivo de configuração com especificações de todos os sensores envolvidos.

4. Conclusão

Sob a ótica do desenvolvedor, este estudo de caso proporciona o conhecimento mais detalhado do *framework* ROS. É possível abstrair o quanto a ferramenta em questão proporciona em termos de funcionalidades para a operação e interação com robôs. A leitura do sensor de toque representa uma pequena demonstração do que é possível fazer utilizando o *framework* ROS e um kit robótico mais simples.

Embora o estudo de caso tenha utilizado um ambiente simples em termos de hardware, a diversidade de plataformas suportadas, o fato de ser livre e de código aberto, sua eficiência quanto a gerência de sistemas robóticos e sua flexibilidade tornam o ROS apto a lidar com desafios em grande escala. Exemplos como o robô PR2 [Cousins, 2010b] mostram o quanto este *framework* é revolucionário e demonstram seu grande potencial quanto ao desenvolvimento de aplicações voltadas a robôs de serviço.

Como trabalho futuro pretende-se utilizar o estudo de caso apresentado e aprofundá-lo considerando o controle dos motores e demais sensores presentes no kit LEGO® Mindstorms® NXT. Com isso, prevalecerá o intuito de criar um primeiro protótipo de robô completo, o qual usa vários sensores e motores em paralelo.

Referências

- Alexander, B., Hsiao, K., Jenkins, C., Suay, B. e Toris, R. (2012). Robot Web Tools. In IEEE Robotics & Automation Magazine, pages 20-23.
- Silva, S. R. X. e Porto, L. (2011) “Análise Comparativa de Kits de Robótica Educativa”, In: XXXIX Congresso Brasileiro de Educação em Engenharia.
- Cousins, S. (2010a). ROS on the PR2. In IEEE Robotics & Automation Magazine, pages 23-25.
- Cousins, S. (2010b). Welcome to ROS Topics. In IEEE Robotics & Automation Magazine, pages 13-14.
- Quirino, G. e Gonçalves, V. (2010) “O Uso da Robótica Educacional no Ensino Fundamental”, In: VIII Encontro Anual de Computação.
- Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R. e Ng, A. (2009) “ROS: an open-source Robot Operating System”, In: International Conference on Robotics and Automation.

Desenvolvimento de um Simulador para o Treinamento de Operações para Profissionais do Setor Elétrico

Felipe G. Santos¹, Iguatemi E. Fonseca²

¹ Pós-Graduação em Sistemas de Comunicação e Automação – UFERSA
Bairro Costa e Silva - Mossoró-RN, CEP: 59.625-900

² Departamento de Sistemas de Computação Centro de Informática – UFPB
Cidade Universitária - João Pessoa-PB, CEP: 58.051-900
santos-fg@hotmail.com, iguatemii@ufersa.edu.br

Abstract. *This Article has the objective of showing a simulator with the educational intent based on operations of the electric sector that will be able to help professionals that realize these tasks constantly as apparel is equipment needed for a particular procedure, tools to use and the procedures to follow before and during the exchange of insulators. And also presents the project and the phases of the development of a simulator highlighting also the architecture use for its implementation.*

Keywords. *Simulator, educational, electric sector, virtual reality.*

Resumo. *Este artigo tem como objetivo mostrar um simulador com o intuito educacional baseado em operações do setor elétrico que possibilitará ajudar profissionais que realizam estas tarefas constantemente, tais como vestimenta de equipamentos necessários para determinado procedimento, ferramentas que deve usar e os procedimentos a seguir antes e durante a troca de isoladores. É apresentado as fases do desenvolvimento do simulador destacando a arquitetura utilizada para sua implementação.*

1. Introdução

Apesar de existir há mais de duas décadas, a Realidade Virtual (RV) tem tido um crescimento ainda maior nos últimos anos por diversos fatores, como preço de equipamentos e grau de conhecimento para utilização de equipamentos que utilizassem a tecnologia.

Com este crescimento o uso das tecnologias da RV em aplicações de treinamento e capacitação de profissionais para a realização de procedimentos complexos ou de risco tem se tornado cada vez mais comum e apresentando resultados muito positivos quando comparado com os métodos tradicionais de ensino, como apostilas, manuais, vídeos entre outros. Com a RV visa-se eliminar barreiras de interação entre o usuário e os sistemas de computador e fornecer uma experiência imersiva convincente Garant (1995).

As vantagens apresentadas pelas soluções em RV são notórias nos dias de hoje, o desenvolvimento de simuladores e aplicações de treinamento em diversas áreas, como na condução de veículos terrestres e aeronaves, manutenção e operação de máquinas

industriais, realização de procedimentos médicos, operações militares e outros. O trabalho apresentado aplica a RV como proposta para o treinamento de operações do setor elétrico utilizando-se de dispositivos computacionais que exploram diversos sentidos humanos, como visão, tato e audição.

2. Fundamentação teórica

A RV é, antes de tudo, uma “interface avançada do usuário” para acessar aplicações executadas no computador, tendo como características a visualização e movimentação em ambientes tridimensionais em tempo real e interação com elementos desse ambiente. Além da visualização em si a experiência do usuário de RV pode ser enriquecida pela estimulação dos demais sentidos como tato e audição Kirner (2007).

Reforçando o conceito, Bowman (2004) diz que a interação do usuário com o ambiente virtual é um dos aspectos importantes da interface e está relacionada com a capacidade do computador detectar e reagir às ações do usuário, promovendo alterações na aplicação.

Tendo esta visão destes conceitos os trabalhos em RV podem pertence a um ou mais grupos, entre eles: i) treinamento de habilidades físicas e motoras Machado (2003); ii) treinamento de tomada de decisão apresentado no trabalho de Souza (2003); permitindo que o usuário execute sequencias reais de montagem de equipamentos; iii) treinamento na execução de procedimentos Nils (2010), o qual consiste de um usuário sem conhecimento de programação de computadores possa criar procedimentos técnicos denominados de instruções técnicas virtuais.

Neste último grupo de aplicações, cuja ênfase no treinamento de procedimentos utilizando técnicas de RV, a ideia principal desta proposta é criar um simulador gráfico de forma a ser semi-imersivo e interativo, reduzindo ao máximo as diferenças existentes entre o cenário virtual e a realidade permitindo aos operadores experimentarem, sem riscos situações de conflito e perigo que ocorrem comumente durante as operações do setor de distribuição de energia.

3. Trabalhos relacionados

Em Hounsell (2007) é apresentado e um modelo de treinamento para manutenção de uma Unidade Hidrelétrica de Energia (UHE) utilizando um sistema de realidade virtual desktop, com uma abordagem de aprendizagem baseada totalmente na prática.

Através do Sistema de Autoria para Construção de Instruções Técnicas Virtuais em Pamplona (2011), permite que técnicos de uma Usina Hidrelétrica de Energia construam instruções técnicas de operação e manutenção em um ambiente virtual dando aos técnicos a possibilidade de interagir com as instruções criadas de modo a montar e desmontar os equipamentos, simulando instruções de manutenção e operação como apresentado na Figura 1.

A Furnas Centrais Elétricas também desenvolveu uma aplicação para o treinamento financiado pela ANEEL (Agencia Nacional de Energia Elétrica), o objetivo central é enriquecer o material utilizado na capacitação das equipes de manutenção, através da Realidade Virtual Belloc (2011).



Figura 1. Fechamento da válvula.

Tendo como alicerce aplicações de sucesso, a ferramenta aqui proposta objetiva tornar a abordagem construtiva mais efetiva ao aprendiz, disponibilizando um ambiente virtual realista com equipamentos e normas de operações executados em sistemas elétricos de distribuição.

4. Desenvolvido do simulador

Nesta seção é apresentada toda a arquitetura do simulador em desenvolvimento que propõe a execução de procedimentos realizados em uma operação de rede elétrica viva (ativa) e morta (desativada). A arquitetura do projeto foi dividido em 3 etapas de forma bem detalhada.

A arquitetura adotada no simulador é baseada em camadas, adaptada para o contexto de jogos e aplicativos para desktop apresentado em Belloc (2011). Nela podemos considerar três camadas essenciais, sendo a camada de interface responsável pela comunicação com o usuário, composta pelas telas do simulador e de menus e sub-menus, a camada de negócio responsável pela lógica/controle do jogo abstraindo dos dados realizados pelo operador, como sequencias de ações a serem realizadas na camada de interface de dados onde temos o ambiente virtual incluindo os objetos de cenas de forma a passar o realismo ao operador como o de uma operação real.

4.1 Apresentação do simulador

Por possuir operações complexas, com risco de vida e de patrimônio tangível e intangível, o setor elétrico necessita de profissionais altamente qualificados e com total conhecimento das ações a serem adotadas para o controle e operação do sistema e segurança. Não se admitindo erros operacionais ou ações precipitadas, pois podem causar perdas de vidas, além de desligamentos indevidos afetando milhares de pessoas e causando prejuízo à empresa.

A Figura 2 mostra a simulação de uma operação que ocorre na troca de um isolador em linha viva de poste trifásica 13,8 kV, à interface interativa controla a comunicação entre o simulador e o operador, caso realize alguma operação incorreta um alerta com sons e uma mensagem é acionada, apresentando qualquer dúvida sobre as normas e recomendações técnicas o operador pode recorrer ao menu de dicas que irá apresentar detalhes e sequências para sua realização da operação.

O simulador apresenta dois modos, demonstrativo e manual, neste o operador vai adquirindo conhecimento durante as etapas da simulação, como quais equipamentos utilizar a roupa para operação, controle da cesta aérea, troca de isoladores, que são acessados conforme sua avaliação no treinamento realizado ao fim de cada etapa,

mostrando a potencialidade e o realismo do uso da RV para o treinamento e avaliação do operador.



Figura 2. Troca de isolador em linha viva.

5. Conclusão e trabalhos futuros

Este artigo apresenta um projeto ainda em desenvolvimento de uma plataforma de treinamento de procedimentos para o setor elétrico tendo inicialmente a execução de duas operações realizadas pelo setor, sendo possível a adição de novas operações (*plugins*) em estudos futuros. Logo após a conclusão pretende-se submeter a plataforma a testes que serão realizados com profissionais do setor e futuramente utilizar a tecnologia do Kinect do console do XBOX 360 para aumentar a imersão da simulação tornando assim a execução dos procedimentos ainda mais realísticos ao operador.

6. Referências

- Belloc, A. R. (2011). Um arcabouço para o Desenvolvimento de Simuladores de Procedimentos em Realidade Virtual. Dissertação de Mestrado. Escola Politécnica da Universidade de São Paulo 2011.
- Bowman, D. A. 3D (2004). “User Interfaces: Theory and Practice”.
- Garant, E. (1995). A virtual reality training system for power-utility personnel. IEEE Pacific Rim Conference on, p. 296 – 299.
- Hounsell, M. P. (2007) Um Ambiente de Treinamento de Manutenção em uma Unidade Hidráulica de Energia Usando Realidade Virtual Desktop. IX Symposium on Virtual and Augmented Reality, p. 286–294.
- Kirner, R. T. C. (2007). Fundamentos de Realidade Virtual e Aumentada. Livro do Pré-Simpósio, VIII Symposium on Virtual Reality, p. 2 – 39.
- Machado, A. C. e L. dos S. (2008). Dispositivos para Sistemas de Realidade Virtual. Realidade Virtual e Aumentada - Uma abordagem tecnológica. Symposium on Virtual and Augmented Reality, Edição Sociedade Brasileira de Computação - SBC, p. 21–45.
- Nils, A. B. P. (2010). Projetos e Implementação de um Sistema de Autoria para Animações, Simulações e Treinamentos em Realidade Virtual. Dissertação de Mestrado. Setembro de 2010. Campus Universitário do Guamá., 2010.
- Pamplona, M. R. F. C. M. A. J. (2003). Um Ambiente de Treinamento de Manutenção em uma Unidade Hidráulica de Energia Usando Realidade Virtual Desktop. 2011.

Disponível em: <http://www.aeseletropaulo.com.br/imprensa/releases/Paginas/AESEletropaulotreinaeletricistasemsimulador3D.aspx> Acesso em jul. 2012.

Souza, M. P. A. (2003). Maintenance and assembly training in a hydroelectric unit of energy virtual reality desktop. IEEE Revista IEEE America Latina, v. 6, n. 5, p. 484–491.

Construção de Software capaz de traçar o Perfil Nutricional e Antropométrico dos estudantes no IFPB campus Campina Grande

Larissa Félix de Andrade¹, Ana Maria da Costa Ribeiro¹, Jossana Pereira de Sousa², Rhavy Maia Guedes¹

¹ Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) – Campina Grande, Paraíba – Brasil

² Departamento de Nutrição, Universidade Federal da Paraíba (UFPB) – João Pessoa, Paraíba – Brasil

{larissafelix.felix,aninhacostaribeiro,jossanasousa,rhavy.maia}@gmail.com

Abstract. *Information technology is increasingly present in areas of human activity. It is possible to build systems that add value to numerous scenarios. One way to encourage nutrition education to the Ministry of Health launched the Food Guide in booklet format evaluated by a nutritionist. However, the capture and analysis of nutritional data for a computerized system can provide reliable results and save time performing activities. Given this context, this work aims at the construction a prototype software capable of profiling nutritional and anthropometric measures of IFPB's students.*

Resumo. *A informática está cada vez mais presente em áreas da atividade humana. É possível construir sistemas que agreguem valor a inúmeros cenários. Como forma de incentivar a educação nutricional o Ministério da Saúde lançou o Guia Alimentar no formato de cartilha avaliado por um Nutricionista. No entanto, a captura e análise dos dados nutricionais por um sistema informatizado pode trazer confiabilidade dos resultados e economia de tempo de realização das atividades. Diante deste contexto, o presente trabalho tem como objetivo a construção de um protótipo de Software capaz de traçar o perfil nutricional e antropométrico dos estudantes do IFPB.*

1. Introdução

Nas ultimas décadas, com a incorporação do avanço científico-tecnológico, a informática está cada vez mais presente nas diversas áreas da atividade humana [Alves et al., 2006], servindo como uma importante ferramenta nos processos de coleta, armazenamento e análise de informações [Stumm et al., 2006]. É possível construir sistemas que agreguem valor às mais diversas áreas do conhecimento. Sua construção geralmente envolve uma equipe multidisciplinar de profissionais, necessários para analisar profundamente todos os aspectos do problema a ser resolvido e aplicar corretamente os preceitos da Engenharia de Software.

A riqueza de informações na área de nutrição constitui fato material a ser empregado no campo da informática, devido às vantagens da utilização de um sistema informatizado, tais como: confiabilidade dos resultados, economia de tempo pela

velocidade dos cálculos, rápida busca de informações, armazenamento do trabalho para consulta posterior, entre outros [Anção, 2001].

Como forma de incentivar a educação nutricional o Ministério da Saúde lançou o Guia Alimentar para a População Brasileira - Promovendo a Alimentação Saudável em formato de panfleto. O objetivo do Guia é orientar os indivíduos a consumir uma alimentação saudável. A aplicação do Guia é acompanhada por um nutricionista que também coleta e analisa informações sobre as medidas antropométricas do paciente.

Diante deste contexto, o presente trabalho tem como objetivo a construção de um protótipo de Software capaz de traçar o perfil nutricional e antropométrico dos estudantes do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB), Campus Campina Grande.

2. Material e Métodos

O protótipo do software foi desenvolvido entre os meses de Fevereiro e Março de 2012. O gerenciamento da equipe técnica foi realizado seguindo uma adaptação do Scrum, combinado ao Extreme Programming (XP) [Schwaber, 2006]. Foram definidas iterações semanais para o desenvolvimento da aplicação. O conjunto de requisitos do protótipo foi elicitados e analisados durante a primeira iteração. Nela foram mantidos contato da equipe técnica com um profissional de nutrição.

2.1 Elicitação, análise de requisitos e viabilidade.

Com o objetivo de melhorar o entendimento das necessidades dos usuários e dos requisitos de negócio foram aplicadas em uma reunião as questões da Tabela 1.

Tabela 1. Elicitação de requisitos para construção do sistema capaz de traçar o perfil dos hábitos alimentares de indivíduos.

Questões	Respostas
Como o processo é feito atualmente?	O Ministério da Saúde por meio do Guia Alimentar de Bolso possui um conjunto de recomendações nutricionais, como também um questionário que verifica hábitos alimentares dos indivíduos. Por meio dele é possível contabilizar pontos, definir padrões de perfil alimentar e sugerir novas condutas. A avaliação antropométrica é realizada a partir da pesagem e medição do indivíduo por um nutricionista. Os dados capturados geralmente são registrados em fichas escritas.
Quais problemas são encontrados nesse processo/ambiente?	Os cálculos realizados para contabilizar as respostas do questionário podem incorrer em erros, pois quando realizados manualmente podem gerar imprecisão, além da possibilidade da aplicação de um número menor de questionários.
Como o sistema ajudaria a reduzir esses problemas?	O sistema calcula automaticamente o IMC e o resultado do questionário de forma rápida e precisa. Desse modo diminui a probabilidade de erro e permite que as atividades sejam realizadas mais rapidamente.
Como o sistema será?	O sistema deve ser capaz de calcular o IMC, aplicar o questionário, sugerir recomendações para novas condutas e armazenar todas as informações inseridas pelos usuários a fim de gerar perfis de hábitos alimentares e antropométricos de uma população.

Após a classificação e organização, os requisitos foram especificados conforme mostra a Tabela 2.

Tabela 2. Lista de requisitos para construção do sistema capaz de traçar o perfil dos hábitos alimentares de indivíduos.

Requisito	Descrição
Calcular o IMC	Por meio do peso, altura e sexo fornecido no cadastro, o cálculo é realizado e possibilita informar se o usuário está acima, abaixo ou em seu estado nutricional adequado ($IMC = PESO/ALTURA^2$).
Avaliar e recomendar hábitos alimentares	Contém questões de caráter avaliatório a respeito do consumo alimentar do usuário. Após a avaliação o usuário receberá orientações para modificar seus hábitos alimentares.
Calcular os gastos calóricos dos indivíduos	A partir do peso, idade, IMC e atividades físicas e diárias será calculado o gasto calórico médio. Nesse procedimento é necessário determinar individualmente as necessidades calóricas e de proteínas, seguindo a recomendação preconizada para o sexo e idade prevista pelas Recomendações de Ingestão Diária (RDA). No entanto para pessoas com idade maior ou igual a 18 anos o cálculo das necessidades calóricas e de proteínas será feito com base nas DRI's - Dietary Reference Intakes.

2.2 Desenvolvimento do sistema

O protótipo do sistema foi desenvolvido na linguagem Pascal, utilizando o paradigma estruturado para a compreensão da sequência de código produzido. A interface de interação com usuário foi construída utilizando controles de tela oferecidos pela linguagem Pascal.

2.3 Avaliação do protótipo

O protótipo foi apresentado aos alunos da disciplina Algoritmos e Lógica da programação que avaliaram suas funcionalidades. Após a utilização aplicou-se as perguntas presentes na Tabela 3.

Tabela 3. Perguntas de Avaliação do protótipo NutriMais.

Questão	Pergunta	Métrica
1	A aplicação oferece uma interação amigável com o usuário?	Sim ou Não
2	O resultado da avaliação foi de claro entendimento?	Sim ou Não
3	É fácil de usar em outro ambiente?	Sim ou Não
4	É fácil entender o conceito e a aplicação?	Sim ou Não
5	É fácil aprender a usar?	Sim ou Não
6	Você gostaria de acessar a aplicação em outra media?	Internet, Dispositivo Móvel.

3. Resultados e Discussão

O resultado da avaliação do protótipo pode ser visto na Tabela 4. Foram entrevistadas 20 pessoas que avaliaram como fácil à interação com a aplicação do Guia Alimentar,

onde os resultados das avaliações são apresentados de forma clara pelo sistema. Além disso, foi sugerida a inserção de imagens para melhor interpretação das etapas de avaliação do Guia e disponibilização do sistema na arquitetura web.

Tabela 4. Resultado da Avaliação

Questão	Métrica	
	Sim	Não
1	19	1
2	20	0
3	18	2
4	18	2
5	20	0

A intenção de migração para plataforma web foi fundamenta pelo resultado da Questão 6 que mostrou o interesse de 13 alunos em acessar esse tipo de mídia e de 7 para acessar a aplicação via dispositivo móvel.

4. Conclusão

A aplicação do protótipo do Guia Alimentar para uma pequena amostra de estudante, e um profissional de nutrição, se mostrou eficaz, mas sua interface não foi aprovada pela maioria deles. Por isso, a arquitetura de desenvolvimento para a próxima iteração do projeto será alterada. O sistema será migrado para a plataforma web que permite maior alcance de usuários via acesso remoto. É importante ressaltar, que o trabalho obteve êxito ao ser submetido ao Comitê de Ética em Pesquisa em Seres Humanos do Instituto Federal de Educação, Ciência e Tecnologia da Paraíba, de acordo com a Resolução no 196/96 do Conselho Nacional de Saúde (Nº041/10).

Referências

- Alves, V. L. S.; Cunha, I. C. K. O.; Marin, H. F. Oliveira O. (2006). Criação de um Web Site para enfermeiros sobre Pé Diabético. 2006.
- Anção, M. S.; Cuppari, L.; Draibe, A. S.; Sigulem D. (2001). Informática em terapia nutricional. In: Magnoni D, Cukier C. Perguntas e respostas em nutrição clínica. São Paulo: Roca. 2001. p. 361-364.
- Stumm, J.; Nassar, S. M.; Pires, M. M. S. (2006). Avaliação alimentar utilizando técnicas de inteligência artificial. In: X CONGRESSO BRASILEIRO EM INFORMÁTICA EM SAÚDE – CBIS, 2006, Florianópolis. Anais CBIS'2006... Florianópolis: CD-ROM.
- Ministério da Saúde. Guia Alimentar – Como Ter uma Alimentação Saudável. (2006). Disponível em: <http://bvsms.saude.gov.br/bvs/publicacoes/guia_alimentar_alimentacao_saudavel.pdf>. Acesso em: 01 jul. 2012.
- Schwaber, K.; Beedle, M. (2001). Agile Software Development with Scrum. 1ª Ed. Prentice Hall, 2001.

Estudo de Caso Usando a Plataforma Microsoft® Robotics Developer Studio (RDS)

Jéssica S. Guimarães, Dalton M. Tavares

Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão – GO – Brasil

{jessica.silva.gui}@gmail.com, {dalton.tavares}@catalao.ufg.br

Abstract. *This paper presents a case study showing the control of a robot via an application developed in C#, in order to demonstrate the use of Microsoft® Robotics Developer Studio (MRDS) for controlling a prototype developed from a LEGO® Mindstorms® NXT kit.*

Resumo. *Este artigo apresenta um estudo de caso apresentando o controle de um robô por meio de um aplicativo desenvolvido em C#, para demonstrar a utilização do Microsoft® Robotics Developer Studio (MRDS) para controlar um protótipo desenvolvido a partir de um kit LEGO® Mindstorms® NXT.*

1. Introdução

Uma forma bastante eficaz para estudos envolvendo programação é a aplicação da mesma em desenvolvimento robótico, colocando em prática os conhecimentos vistos em cursos de graduação e programando protótipos de robôs, por meio de softwares existentes. Estes protótipos podem ser construídos a partir de kits de desenvolvimento disponíveis em mercado, de fácil manuseio e custo acessível.

Com os avanços tecnológicos na área da robótica, surgiram diversos kits de desenvolvimento como o robô NAO da Aldebaran Robotics¹, o Bioloid² e o LEGO® Mindstorms^{®3}, entre outros. Para o presente estudo de caso, o kit desenvolvido pela LEGO® Mindstorms® foi escolhido, devido a sua flexibilidade e baixo custo frente às demais opções.

Em 2006, a LEGO® lançou uma nova linha de produtos, o LEGO® Mindstorms® NXT, incentivando a educação tecnológica de forma descontruída [Kim e Jeon 2008]. O kit de desenvolvimento é constituído por peças para montagem rápida, incluindo um bloco de controle programável, sensores de toque, para medição de intensidade luminosa, ultrassônico, motores, entre outros. Seu ambiente de programação nativo permite a aplicação de comportamentos que utilizam as ferramentas que o acompanham através da programação icônica⁴.

O bloco de controle pode ser programado usando outros softwares, diferentes do ambiente de programação padrão fornecido pela LEGO®. Uma das opções é o Microsoft® Robotics Developer Studio (MRDS), aplicável ao controle e simulação de robôs, destinado

¹ Para maiores informações, vide: <http://www.aldebaran-robotics.com/en/> Acesso em: 28/01/2013.

² Para maiores informações, vide: http://robosavvy.com/store/product_info.php/products_id/82 Acesso em: 28/01/2013.

³ Para maiores informações, vide: <http://mindstorms.lego.com/en-us/Default.aspx> Acesso em: 28/01/2013.

⁴ Uma forma simples de programar, utilizando metodologia gráfica, ao invés de código fonte inserido como texto, permitindo a construção de cenários e a interpretação visual.

a desenvolvedores acadêmicos, amadores e comerciais.

Este artigo tem por objetivo apresentar um estudo de caso envolvendo o controle de um robô móvel, desenvolvido usando o kit LEGO® Mindstorms® NXT e o software RDS. Para tanto, será fornecida uma breve descrição das principais funcionalidades do software RDS (seção 2), seguida da apresentação do ambiente de desenvolvimento e dos resultados alcançados (seção 3). Uma breve conclusão e propostas de trabalhos futuros serão apresentadas na seção 4.

2. Estrutura do RDS

A arquitetura básica do RDS é apresentada na Figura 1. De forma sucinta, o RDS fornece serviços de alto nível trabalhando normalmente com abstrações de serviços de baixo nível, as quais, por sua vez, tratam da comunicação com dispositivos. Os serviços de baixo nível são fornecidos pelos fabricantes de robôs ou desenvolvidos pelo próprio usuário, por meio de aplicativos escritos em C#, por exemplo. É através de sua camada de abstração de hardware, com interface bem definida, que os serviços de alto nível têm controle de entrada e saída (I/O). Os serviços CCR e DSS (explicados abaixo) são componentes característicos do RDS e atuam no contexto de comunicação entre nós. Esse modelo é baseado no *framework* .NET. Esse conjunto de serviços se comunica com o *firmware* do robô através de uma porta serial – USB - ou conexão via rede, através do Windows, o qual é fundamental para que o serviço de interação do robô com o aplicativo seja realizado [Microsoft Robotics Group 2012].

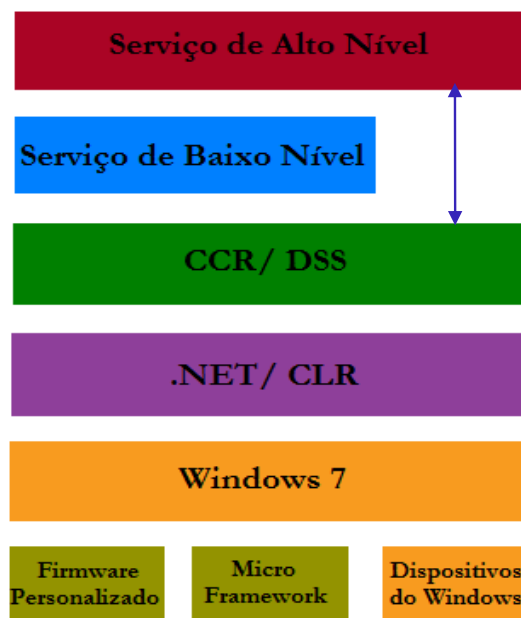


Figura 1. Arquitetura do software.

A Concorrência e Coordenação Runtime (CCR – *Concurrency and Coordination Runtime*) é um modelo que trabalha com programação multi-thread e sincronização de

tarefa, e permite interação com robôs em tempo real sem a complexidade de trabalhar com semáforos, exclusão mútua, tratamento de possíveis assincronias, entre outras, devido o uso de protocolo aberto. Já os Serviços de Software Distribuídos (DSS - *Distributed Software Services*) é um programa que atuam juntamente com CCR, utilizando vários serviços ou instancias de serviços, tais como: componentes de hardware (sensores e atuadores), componentes de software (repositório, interface de usuário), fusão entre sensores e tarefas relacionadas aos mesmos. [Cepeda, Chaimowicz e Soto 2010].

3. Estudo de caso: utilizando sensor de toque para encontrar obstáculos

Utilizando um kit LEGO® Mindstorms® NXT, foi estruturado um protótipo que tem como principais características o uso de dois motores, para acionar as rodas, o uso de um sensor de toque, para reconhecer os obstáculos que lhe forem impostos através de contato direto e um bloco de controle NXT. Utilizando o RDS foi desenvolvido um aplicativo que controla o movimento dos motores do protótipo fazendo com que ele se movimente no espaço e, ao identificar um toque por meio do sensor, emite uma mensagem informando que encontrou um obstáculo. Este aplicativo foi desenvolvido na linguagem C#, usando o Microsoft® Visual C# Express, um dos softwares utilizados na plataforma RDS. A Figura 2 mostra um exemplo de robô utilizando um sensor de toque e motores para se movimentar.

Para utilizar a plataforma RDS, primeiramente instalamos o sistema operacional Windows 7, o Microsoft® Visual C# Express, o Kinect para Windows SDK, o qual pode ser utilizado futuramente em projetos envolvendo o reconhecimento de movimentos por meio de um sensor Microsoft® Kinect⁵ e o RDS em sua última versão⁶.



Figura 2. Protótipo indo em direção a uma barreira.

Após a preparação do ambiente de desenvolvimento, a comunicação entre o bloco de controle e o computador foi configurada, com o intuito de permitir o controle do protótipo utilizando o programa desenvolvido na plataforma RDS. De forma sucinta, foi criada uma conexão do notebook com o bloco de controle via *Bluetooth*, adicionando o bloco de controle como dispositivo do notebook, de forma que ambos se reconhecessem e

⁵ Para maiores informações, vide: <http://www.microsoft.com/en-us/kinectforwindows/>

⁶ Até a data de submissão do artigo, o software RDS estava em sua versão 4.0. Fonte: <http://www.microsoft.com/robotics/>. Acesso em: 11/01/2013.

pudessem trocar informações.

Para obter um melhor desempenho é indicado que se faça uma atualização da versão do *firmware* do bloco de controle para uma versão mais recente⁷. Esta atualização é feita usando o software que é fornecido juntamente com o kit LEGO® Mindstorms® NXT, instalado em um computador⁸.

3.1. Controle usando C#

Usando o Visual C# Express, foi desenvolvido um aplicativo, o qual imprime uma mensagem ao colidir com um obstáculo durante seu percurso. O primeiro passo é criar um novo serviço utilizando o interpretador de comandos DSS, o qual fornece suporte ao protocolo descentralizado de serviços de software (Decentralized Software Services Protocol – DSSP). No Visual Studio é necessário adicionar uma referência à biblioteca `RoboticsCommon.Proxy.dll`, a qual fornece suporte ao sensor de toque. Observe que ao iniciar o aplicativo (neste caso chamado de `MyTutorial1`) é gerado um manifesto que aparece no lado direito da tela (Figura 3 e 4), e diz respeito as configurações de depuração fornecidas pelo `DSSNewService`.

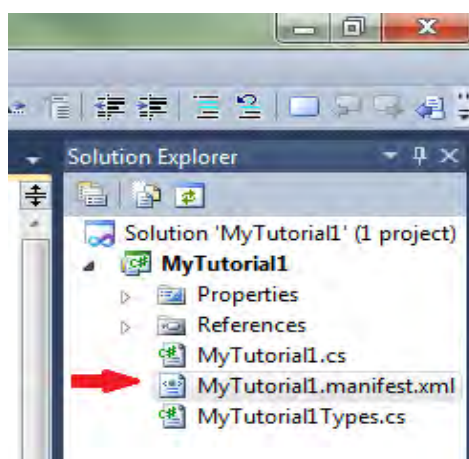


Figura 3. Manifesto criado na inicialização do aplicativo.

⁷ A versão usada foi 1.29, entretanto existe uma versão mais recente – a 1.31. Fonte: <http://mindstorms.lego.com/en-us/support/files/default.aspx#Firmware> Acesso em: 11/01/2013.

⁸ Usando a opção *update*, conecta-se o bloco de controle à porta USB do computador e se faz o download do novo *firmware*. Fonte: <http://msdn.microsoft.com/en-us/library/bb483050.aspx> Acesso em: 28/01/2013.

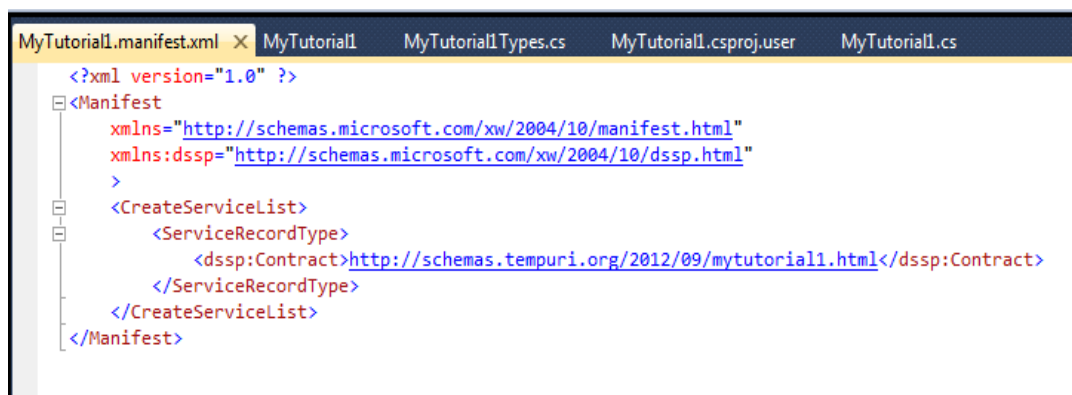


Figura 4. Descrição do manifesto.

É necessário modificar o manifesto de acordo com o tipo de hardware usado no âmbito da guia Debug. Neste caso, a configuração ficou conforme a descrição fornecida na Listagem 1.

```
/p:50000 /t:50001
/m:"samples/MyTutorial1/MyTutorial1.manifest.xml"
/m:"samples/config/LEGO.NXT.MotorTouchSensor.manifest.xml"
```

Listagem 1. Referência ao manifesto adequado ao nosso hardware.

Feito isso, um serviço de comunicação é criado e inicializado entre o aplicativo e o sensor de toque. Ao final do código insere-se a assinatura do código que permitirá a execução do serviço de notificação, o qual recebe a notificação de que o sensor foi pressionado e então chama o método `BumperHandler()`. Este método envia a mensagem para o console, de forma que a informação possa ser visualizada. Na Listagem 2 são apresentadas as linhas que foram inseridas no código.

```
//Criando o serviço de comunicação entre este aplicativo e o sensor de
toque
[Partner("bumper", Contract = bumper.Contract.Identifier,
  CreationPolicy = PartnerCreationPolicy.UseExisting)]
private bumper.ContactSensorArrayOperations _bumperPort = new
bumper.ContactSensorArrayOperations();

// Iniciando serviço que ouve o bumper.
SubscribeToBumpers();
}

/// <summary>
/// Subscrevendo serviço de bumper.
/// </summary>
void SubscribeToBumpers()
{
```

```

        // Criando porta de notificação do bumper.
        bumper.ContactSensorArrayOperations bumperNotificationPort = new
bumper.ContactSensorArrayOperations();

        // Subscrive o serço bumper e recebe notificações no
bumperNotificationPort.
        _bumperPort.Subscribe(bumperNotificationPort);

        // Recebe atualização das notificações do bumperNotificationPort .
        Activate(
            Arbiter.Receive<bumper.Update>
                (true, bumperNotificationPort, BumperHandler));
    }

    /// <summary>
    /// Recebe a notificação de que o sensor foi pressionado e envia a
informação para o console
    /// </summary>
    /// <param name="notification">Atualiza notificação</param>
    private void BumperHandler(bumper.Update notification)
    {
        if (notification.Body.Pressed)
            LogInfo(LogGroups.Console, "Ops, um obstáculo foi encontrado!");
    }

```

Listagem 2. Código fonte do aplicativo desenvolvido.

Ao executar o programa, é aberta uma janela no navegador com o endereço padrão <http://127.0.0.1:50000/>, que mostra as configurações do bloco de controle, do sensor e informações referentes à comunicação entre o aplicativo e o protótipo (Figura 5).

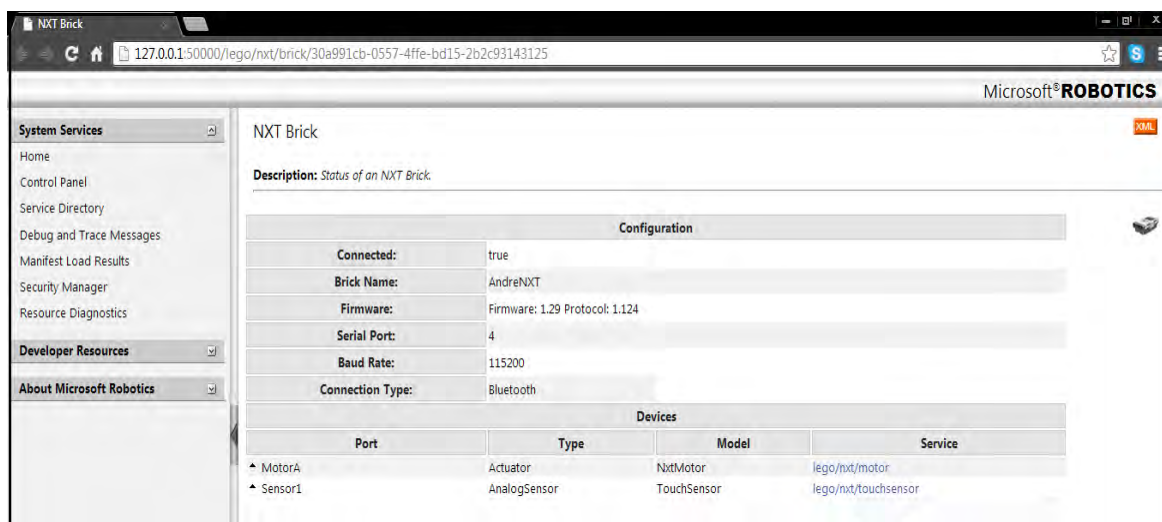


Figura 5. Página inicial do console de saída do serviço.

Clicando em Debug e mensagens de rastreamento, à esquerda da página podemos visualizar detalhes do serviço, como detalhes das mensagens, detalhes da comunicação com o hardware, erros, a mensagem de que o sensor foi pressionado, entre outros (Figuras 6 e 7).

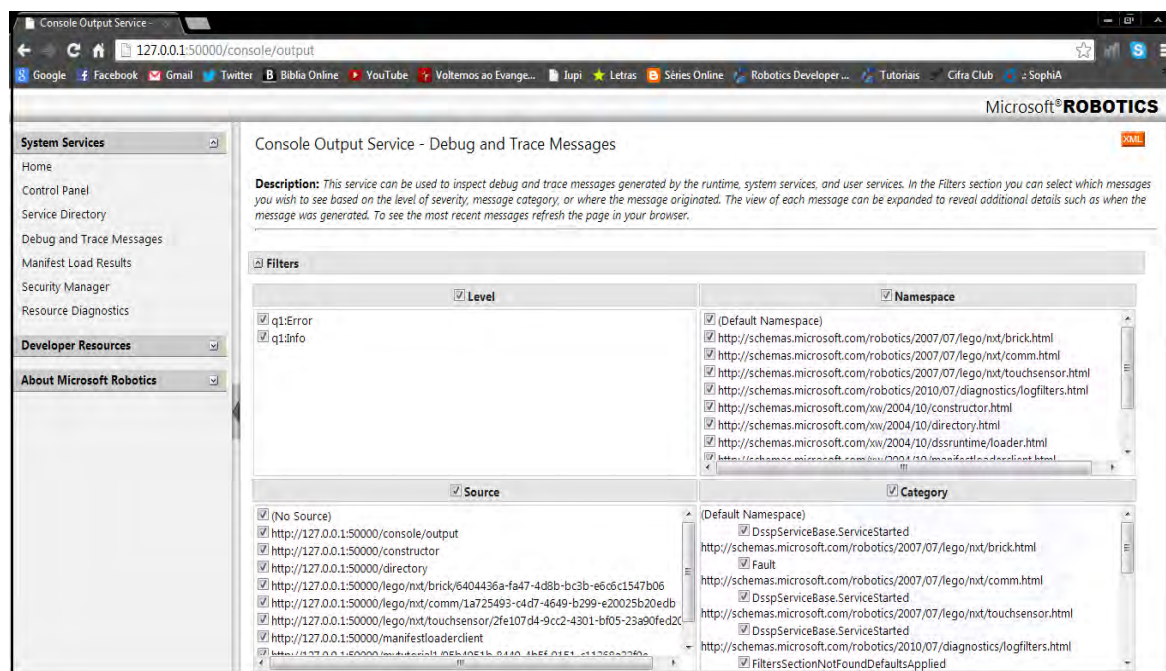


Figura 6. Console de saída do serviço – Debug e rastreamento de mensagens.

16	12:52:57	*	Mounted embedded resources for assembly 'file:///C:/Users/Jhessyka/Microsoft Robotics Dev Studio 4/bin/roboticscommon.dll' under prefix '/resources/RoboticsCommon/'
17	12:52:57	*	Mounted embedded resources for assembly 'file:///C:/Users/Jhessyka/Microsoft Robotics Dev Studio 4/bin/mytutorial1.y2012.m09.dll' under prefix '/resources/MyTutorial1.Y2012.M09/'
18	12:53:16	*	Ops, um obstáculo foi encontrado! Namespace http://schemas.microsoft.com/xw/2005/12/dss.html Category Console Level q1:Info Time 2013-01-30T12:53:16.1324631Z Subject Ops, um obstáculo foi encontrado! Source http://127.0.0.1:50000/mytutorial1/cf0eb356-8de7-4a4e-8ebd-bdf160b3081b CodeSite Void BumperHandler(Microsoft.Robotics.Services.ContactSensor.Proxy.Update) at MyTutorial1.MyTutorial1Service System.Collections.Generic.IEnumerator`1[Microsoft.Ccr.Core.ITask] Execute() at Microsoft.Ccr.Core.Task`1[T0] Microsoft.Ccr.Core.IteratorContext ExecuteTaskHelper(Microsoft.Ccr.Core.ITask) at Microsoft.Ccr.Core.TaskExecutionWorker Void ExecuteTask(Microsoft.Ccr.Core.ITask ByRef, Microsoft.Ccr.Core.DispatcherQueue, Boolean) at Microsoft.Ccr.Core.TaskExecutionWorker Void ExecutionLoop() at Microsoft.Ccr.Core.TaskExecutionWorker Void Run(System.Threading.ExecutionContext, System.Threading.ContextCallback, System.Object, Boolean) at System.Threading.ExecutionContext Void Run(System.Threading.ExecutionContext, System.Threading.ContextCallback, System.Object) at System.Threading.ExecutionContext Void ThreadStart() at System.Threading.ThreadHelper

Figura 7. Mensagem apresentada quando o sensor é pressionado.

4. Conclusão

Depois de preparar o computador, configurar o bloco de controle NXT e confeccionar o aplicativo, observa-se a interação do *framework* RDS com o bloco de controle. A mensagem referente à detecção de barreira foi impressa em uma página do navegador (Figura 7), a qual também contém informações detalhadas sobre o protótipo (Figura 5), detalhes da comunicação com o hardware, entre outros (Figura 6).

Com este experimento, utilizando o Visual Basic é possível controlar um protótipo desenvolvido a partir do kit LEGO® Mindstorms® NXT e aplicativos escritos em C#. Também é possível observar a existência de uma boa interação entre o *software* RDS e o bloco de controle, o que permite a realização de uma ampla gama de experimentos.

Um ponto negativo observado durante a pesquisa é que não há simultaneidade com relação à detecção de toque e a impressão da mensagem na tela, sendo necessário atualizar a página aberta pelo programa para visualizar as informações atuais.

Como trabalho futuro, podemos realizar o controle de protótipos usando a interatividade com o sensor Microsoft® Kinect, capaz de captar movimentos através de rastreamento do corpo humano, o que pode permitir futuras implementações em escala maior, como por exemplo, no controle de braços mecânicos, os quais são amplamente utilizados pelas indústrias, como a automobilística, entre outras [Suay e Chernova 2011].

Projetos como o SMERobot⁹, desenvolvido pela União Europeia, propuseram a aplicação de tecnologias inovadoras aplicáveis ao controle de robôs de modo a tornar a interação com seus programadores e operadores mais intuitiva. Uma aplicação sugerida e bastante promissora neste ramo é o reconhecimento de gestos e voz, o qual permite uma fácil interatividade do usuário com o robô [Neto e Pires 2009]. Ao invés de descrição de serviços utilizando linguagens de programação, os operadores mostrariam o serviço que deve ser feito por meio de gestos. Esses dados seriam captados por sensores, armazenados e enviados ao robô através de *Bluetooth*.

Referências

- Kim, S. H. e Jeon, J. W. (2008) “Using visual programming kit and Lego Mindstorms: An introduction to embedded system”, In: Proceedings of the IEEE International Conference on Industrial Technology (ICIT).
- Neto, P. e Pires, J.N. (2009) “High-Level Programming for Industrial Robotics: using Gestures, Speech and Force Control”. Submitted to the IEEE International Conference on Robotics and Automation (ICRA2009), Kobe, Japan, 2009.
- Suay, H.B. e Chernova, S. (2011) “Humanoid robot control using depth camera”, In: 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI).
- Cepeda J.S., Chaimowicz L. e Soto R. (2010) “Exploring Microsoft Robotics Studio as a Mechanism for Service-Oriented Robotics”, In: 2010 Latin American Robotics Symposium and Intelligent Robotics Meeting.
- Microsoft Robotics Group (2012) “Robotics Developer Studio: Reference Platform Design v1.0”, In: Documentation of the Microsoft Robotics Developer Studio 4.

⁹ Maiores informações em: <http://www.smerobot.org/> Acesso em: 28/01/2013.

Estudo de Caso Usando o Framework LeJOS

Nilton M. de Souza¹, Dalton M. Tavares¹

Departamento Ciências da Computação – Universidade Federal de Goiás
Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão – GO –
Brasil

nilton.mendes.2@gmail.com, dalton.tavares@catalao.ufg.br

Abstract. *This paper describes how to use the LeJOS framework in the scope of an experience that deploys a route, using elements called WayPoints. WayPoints determine “points” in space, previously marked by the user, describing a path to be followed. The LeJOS is an open source project that allows the use of the Java language to control the LEGO® Mindstorms® NXT kit. It has a powerful API that allows the easy implementation of activities, using sensors and motors. The project uses a simple mobile robot, made from the LEGO® kit, showing how the logic and coding was made in order to reach the objective.*

Resumo. *Este artigo descreve a utilização do framework LeJOS no âmbito de um experimento que desenvolve uma rota, através de elementos chamados WayPoints. WayPoints determinam “pontos” no espaço, previamente marcados pelo usuário, formando uma trajetória a ser seguida. O LeJOS é um projeto open source que permite a utilização da linguagem Java para manipulação do kit LEGO® Mindstorms® NXT. O LeJOS possui uma poderosa API que permite implementar várias atividades facilmente, como manipulação de sensores e motores. O projeto utiliza um robô móvel simples, que é feito a partir do kit LEGO®, mostrando como foi feita a programação e a lógica utilizada para alcançar o objetivo.*

1. Introdução

LeJOS é um *framework* para desenvolvimento de aplicações usando o kit LEGO® Mindstorms® NXT. Para desenvolver este estudo de caso foi usado um kit LEGO® Education modelo 9797, o qual contempla um bloco inteligente, rodas, pneus, engrenagens e peças para montagem de protótipos, motores e sensores. Neste caso, o kit foi usado para montar um pequeno robô móvel, a ser utilizado no experimento. As peças fundamentais para a montagem de qualquer protótipo são três motores, quatro sensores (som, ultra-sônico, intensidade luminosa e toque), o bloco inteligente com uma pequena tela de LCD, cabos para conectar os motores e sensores ao bloco inteligente, um cabo USB para conectar o bloco inteligente ao computador e interface *Bluetooth*. [Oliveira, Thiago et al. 2008]

O *framework* LeJOS proporciona a utilização da linguagem JAVA, com várias bibliotecas e recursos distribuídos, em conjunto com uma máquina virtual bem simples, a qual é inserida no bloco inteligente substituindo o *firmware* original fornecido pela LEGO®. Além disso, também existem algumas bibliotecas para a contra parte hospedada em um computador, incluindo vários exemplos de programas para a operação do robô

ou dispositivo, controlado via computador ou via programa hospedado no bloco inteligente NXT [Lew 2010].

O LeJOS detém uma API que possui bibliotecas que fornecem suporte ao sensor luminoso, sensor ultra-sônico, sensor de toque, sensor sonoro, aos três motores e o estabelecimento de conexão via USB ou Bluetooth. Outro recurso importante é o suporte a *multithreading*, muitas vezes necessário para que duas ou mais atividades concorrentes sejam realizadas.[LeJOS Team 2008c] Um exemplo seria o uso concorrente de dois sensores e dois motores, onde cada um compõe uma tarefa diferente, sendo representados cada um em uma *thread* específica.

Conexões via USB e *Bluetooth* são muito importantes, pois permitem a expansão dos recursos nativos do bloco inteligente. Este é muito limitado comparado a um computador ou *smartphones* atuais. Dessa forma, parte do processamento pode ser realizado fora do bloco inteligente (ex. em um notebook ou PC) e informações de interesse podem ser trocadas via USB ou *Bluetooth*, garantindo o controle do dispositivo pretendido. Essas conexões também são utilizadas para enviar os programas ao bloco inteligente.

Utilizando o *framework* LeJOS e o kit de desenvolvimento LEGO® Mindstorms® NXT, este artigo irá apresentar um estudo de caso envolvendo um planejador de trajetória simples por meio da determinação manual de pontos. Para tanto, será apresentada uma breve revisão bibliográfica sobre o *framework* LeJOS na seção 2, o estudo de caso será descrito na seção 3 e uma breve conclusão e propostas de trabalhos futuros serão discutidos na seção 4.

2. Organização do Framework LeJOS

O LeJOS permite a programação de robôs LEGO® Mindstorms® NXT/RXC utilizando a linguagem JAVA. Este *framework* surgiu de uma evolução do TinyVM, que é uma pequena máquina virtual para Java e um *firmware* para substituir o sistema original do LEGO® Mindstorms® NXT [LeJOS Team 2008b]. .

Para instalação, o próprio site do LeJOS fornece o *link* para *download* do pacote, através de um repositório no site *sourceforge*. Este código está disponível para Mac OS, Linux e Windows necessitando apenas de alguns pacotes descritos na seção 3.

O pacote contém:

- Um *firmware* para substituir o *firmware* original presete no bloco inteligente do kit LEGO® Mindstorms®;
- Uma biblioteca que auxilia no desenvolvimento de aplicativos para PC e para o bloco inteligente NXT;
- Um endereçador (*linker*) que transforma o código java, para uma linguagem compreendida pelo bloco inteligente do NXT;
- Ferramentas para compilação, envio de arquivos, modificação do *firmware* e *stream* entre o computador e o NXT;
- Vários exemplos de programas e códigos .

3. Estudo de Caso

O LeJOS possui uma vasta documentação para programação em sua linguagem, instalação, configuração, além de tutoriais iniciais¹. Para o contexto desse artigo, será usada a versão para Linux Ubuntu 12.04 em conjunto com a IDE Eclipse® configurada para o desenvolvimento em JAVA e um *plugin* específico para o LeJOS. Essa configuração facilita a criação, compilação e envio dos programas para o bloco inteligente.

Além da máquina virtual Java e do software LeJOS, também deverão ser instalados o “ant” e a biblioteca “lib-usb”. O ant² é um *maker*, baseado em Java, seu objetivo é construir projetos em Java a fim de serem executados em qualquer máquina que tenha uma máquina virtual java. A lib-usb é importante pois permite uma conexão via *usb* entre dispositivos sem que exista um *driver* instalado.

Após a instalação do ambiente de desenvolvimento, deve ser criado um usuário (ex. *lego*) e a regra de operação do dispositivo para o gerenciador de dispositivos (*udev*). A regra criada permite a descoberta do bloco inteligente via barramento USB e a definição do grupo de acesso e permissões (Listagem 1).

```
BUS=="usb", SYSFS{idVendor}=="03eb", GROUP="lego", MODE="0660"
BUS=="usb", SYSFS{idVendor}=="0694", GROUP="lego", MODE="0660"
```

Listagem 1. Regras para o udev definidas em /etc/udev/rules.d/lego-mindstorms-nxt.rules.

O estudo de caso proposto visa apresentar um planejador de trajetória simples. Em uma implementação posterior, pode ser realizado o planejamento de movimento (*motion planning*). O planejamento de movimento leva em consideração duas tarefas bem definidas: planeja o caminho (*path planning*) e planeja a trajetória (*trajectory planning*). O planejamento de caminho diz respeito a geração de um caminho livre de colisões em um ambiente com obstáculos e sua otimização com relação a algum critério. Um planejador de trajetória deve fornecer movimentos pré-programados ao longo de um caminho planejado [Hachour 2008].

O planejador de trajetória pretendido deve permitir que seja traçada uma rota manualmente, inserindo pontos em um padrão de *waypoints* dados três pontos X, Y e uma orientação (ou *heading*) que consiste na posição em relação a frente do carro, evitando que ele percorra um dado grau de rotação pelo pior caminho, o que poderia prejudicar a execução do caminho traçado. Posteriormente a rota armazenada deve ser executada a partir dos pontos gravados e suas orientações. Uma representação do protótipo do veículo é apresentada na Figura 1.

O protótipo construído possui duas rodas frontais, e dois motores conectados a cada uma, e uma pequena roda traseira. Os motores são conectados por meio de cabos RJ11 ao bloco inteligente que os controlam, e as rodas são ligadas por pequenas engrenagens. A estrutura do veículo é feita de pequenos componentes plásticos do kit.

1 Para maiores informações acesse <http://lejos.sourceforge.net/>. Acesso em: 14/01/2013.

2 Para maiores informações acesse <http://ant.apache.org/manual/>. Acessado em: 17/01/2013

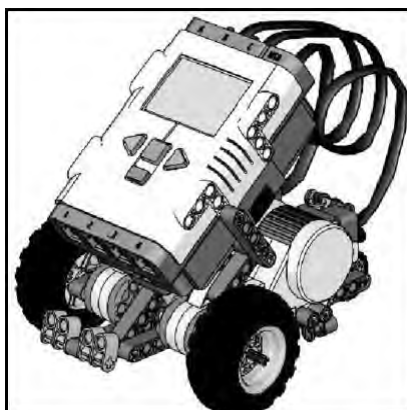


Figura 1: Protótipo do veículo

A seção 3.1 fornece as bases para o experimento proposto usando o *framework* LeJOS. Para tanto, serão descritos os fundamentos de operação dos motores e as funções utilizadas para o planejamento de trajetória.

3.1. Planejamento de Trajetória no LeJOS

Para se desenvolver um planejador de trajetória simples usando o LeJOS, deve-se levar em conta a leitura dos sensores e motores do kit LEGO® Mindstorms® NXT. Estes são diretamente conectados a portas específicas, sendo os motores ligados às portas, A, B e C e os sensores às portas 1, 2, 3 e 4. Não existe uma especificação formal de uma dada ordem para a ligação de motores ou sensores, porém, o programador deve saber a qual porta um dado dispositivo está relacionado. Para o estudo de caso em questão, serão apresentados os fundamentos de operação dos motores utilizando o *framework* LeJOS.

3.1.1. Manipulando Motores

Conforme discutido anteriormente, os motores do kit LEGO® Mindstorms® NXT podem estar associados a três portas: A, B ou C. As classes utilizadas para manipular os motores, na maioria dos casos, são a `lejos.nxt.NXTRegulatedMotor` e a `lejos.nxt.Motor`.

`NXTRegulatedMotor` é a abstração de um motor e `Motor` possui três instâncias de `NXTRegulatedMotor` (A, B ou C), cada uma relacionada a uma porta destinada a um motor e o acesso a cada uma é dado por `Motor.A`, `Motor.B` e `Motor.C`. Algumas das propriedades previstas para os motores são:

- `forward()`: rotaciona o motor para frente, até que seja invocada a parada mediante o uso de `flt()` ou `stop()`;
- `backward()`: Rotaciona o motor para trás, até que seja invocada a parada mediante o uso de `flt()` ou `stop()`;
- `stop()`: para o motor;
- `flt()` ou `flt(boolean imediato)`: para o motor e o desliga, colocando-o em modo flutuante. No segundo caso, o argumento indica se a instrução deve ou não ser executada imediatamente;
- `getAcceleration()`: retorna a aceleração atual do motor;

- `getPosition()`: retorna a posição que o motor está tentando manter;
- `getTachoCount()`: fornece a contagem atual de rotação do motor (valor do tacômetro);
- `getSpeed()`: retorna a velocidade do motor;
- `resetTachoCount()`: zera a contagem de rotação do motor (tacômetro);
- `rotateTo(int angulo)`: movimenta o motor até dado ângulo a partir de 0;
- `setAcceleration(int aceleracao)`: define a aceleração do motor;
- `setSpeed()`: define a velocidade do motor.

A algumas operações que podem ser realizadas com motores, como por exemplo, a rotação para frente, para trás, a verificação da aceleração e o ajuste de velocidade. Outra questão de suma importância para o controle de um robô, de um modo geral, diz respeito a sua operação remota. Dessa forma, é possível conectar o robô por intermédio do bloco inteligente com o computador, via *Bluetooth* e controlá-lo usando uma interface simples [Oliveira, Geraldo 2009].

O controle remoto dos motores (Listagem 1) é realizado usando uma classe `Motor` semelhante ao exemplo mostrado na Listagem 2. Cada motor é do tipo `NXTRemoteMotor`, o qual é semelhante a `NXTRegulatedMotor`, entretanto, este é utilizado no computador para controlar remotamente os motores A, B e C.

```
import lejos.nxt.Motor;

public class TesteMotor { //Operações com motores porta A
    public static void main(String[] args) throws Exception {
        Motor.A.forward();//Rotação para frente
        Motor.A.flt();//Para o motor, coloca em modo flutuante
    }
}
```

Listagem 1 . Demonstração de código para controle remoto de motores

3.1.2 Planejamento de Trajetória no LeJOS

A implementação utilizou o conceito de *waypoints* no LeJOS (Listagem 3). *Waypoints* são marcações realizadas nos eixos X, Y e dotadas de um fator H, em graus, o qual determina uma orientação em relação a frente do veículo simulado. Esta orientação refere-se a informação pré-determinada sobre as “curvas”, para que o robô não se perca ou se movimente em graus excessivos [LeJOS Team 2008a].

Para desenvolver essa atividade foi instanciado um piloto do tipo `DifferentialPilot` que é a abstração de um carro com duas rodas, e dois motores. Este precisa apenas de informação quanto a dimensão das rodas e quais motores serão utilizados à direita e esquerda. Essa classe já implementa diversas tarefas que um carro faz, como andar, virar, parar entre outras.

Também fez-se o uso da classe `OdometryPoseProvider`, a qual é responsável por designar a posição para o carro. Ela também funciona como um *listener* de

movimento. Ao final, foi usada a classe `Navigator` que controla o robô através de um caminho (*path*), em uma sequência de *waypoints*, até que alcance o último ponto e pare. Para construir o *path* são adicionados pontos através do método `addWaypoint(new Waypoint(X, Y, H))`. Em que os pontos X e Y são pontos de um plano cartesiano que determina sua localização. Exemplo de como funciona o plano cartesiano na Figura 2.

Como este exemplo é bem simples, os pontos de *heading* foram lançados manualmente, mas o cálculo para encontrar esse valor é dado pela posição do *waypoint* atual até o seguinte, formando uma reta. Esta reta forma um ângulo com a atual posição da frente do robô, até estar na direção desta reta, que é a mesma direção do próximo ponto. Para obter o ângulo utilizam-se a equação do coeficiente angular da reta (Equação 1), que dá o valor do coeficiente da reta em relação a tangente e a equação para o cálculo do arco da tangente (Equação 2).

$$tgz = \frac{y_B - y_A}{x_B - x_A}$$

Equação 1. Cálculo do coeficiente angular da reta

$$\arctan z = \sum_{n=0}^{\infty} \frac{2^{2n} (n!)^2}{(2n+1)!} \frac{z^{2n+1}}{(1+z^2)^{n+1}}$$

Equação 2. Fórmula do arco tangente

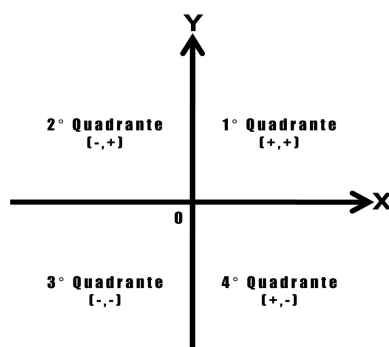


Figura 2: Plano Cartesiano

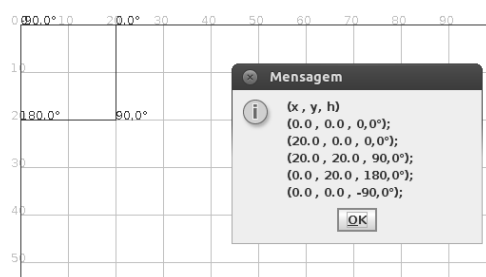


Figura 3: Demonstrativo de pontos em uma trajetória


```

import lejos.nxt.Button;
import lejos.nxt.Motor;
import lejos.nxt.LCD;
import lejos.robotics.navigation.DifferentialPilot;
import lejos.robotics.navigation.Navigator;
import lejos.robotics.navigation.Waypoint;
import lejos.robotics.navigation.Pose;
import lejos.robotics.pathfinding.Path;
import lejos.robotics.localization.OdometryPoseProvider;
public class NavigatorTest {
    Navigator nav;
    public static void main(String[] args) throws IOException {
        System.out.println("Pressione um Botão");
        Button.waitForAnyPress();
        DifferentialPilot pilot = new DifferentialPilot(6.5f,
            11.5f, Motor.A, Motor.B);
        OdometryPoseProvider pp = new OdometryPoseProvider(pilot);
        pilot.addMoveListener(pp);
        Navigator nav = new Navigator(pilot);
        //Adiciona waypoints (X, Y, H)
        nav.addWaypoint(new Waypoint(0, 0, 0));
        nav.addWaypoint(new Waypoint(20.0, 0.0, 90));
        nav.addWaypoint(new Waypoint(20.0, 20.0, 180));
        nav.addWaypoint(new Waypoint(0, 20, -90));
        nav.addWaypoint(new Waypoint(0, 0, 0));
        //Segue os pontos marcados acima.
        nav.followPath();
        Button.waitForAnyPress();
        nav.followPath();
        System.out.println(nav.getPath());
        Button.waitForAnyPress();
    }
}

```

Listagem 2. Implementação do planejador de trajetória usando waypoints

Este será o valor do *heading* para este trajeto entre os dois pontos. Em Java pode-se usar o método $\text{Math.atan2}(y_B - y_A, x_B - x_A)$, o qual retorna o grau de inclinação da reta. Um exemplo gráfico é apresentado na Figura 3.

4. Conclusão

A utilização da ferramenta LeJOS para o desenvolvimento desta pesquisa foi adequada, dada a sua vasta documentação e um fórum bastante ativo. O problema a ser resolvido não é um problema complexo, considerando que a API LeJOS implementa as funções necessárias, facilitando o processo de desenvolvimento, bastando apenas utilizar a fundamentação matemáticas para finalizar a pesquisa.

Como proposta de trabalho futuro, pretende-se criar um planejador de caminho como parte de um estudo para o desenvolvimento de um planejador de movimento completo, o qual terá uma contra parte no PC e uma no NXT. No PC será implementada uma tela semelhante a apresentada na Figura 2, e ao final da marcação dos pontos, estes serão enviados para o NXT, o qual deve criar o *path* a ser seguido. O experimento deve usar sensores luminosos e ultra-sônico, pois o caminho traçado poderá ter obstáculos. Dessa forma, o veículo estará apto a prever o risco de colisão e desviar, sem perder sua localização, deverá retornar a trajetória assim que possível.

Referências

- Lew, M.W., Horton, T.B. e Sherriff, M.S. (2010) “Using LEGO MINDSTORMS NXT and LEJOS in an Advanced Software Engineering Course”, In: *23rd IEEE Conference on Software Engineering Education and Training (CSEE&T)*, pp.121-128.
- LeJOS Team (2008a) “LeJOS NXJ API”, Disponível em: <http://lejos.sourceforge.net/nxt/nxj/api/index.html>. Acessado em: 23/01/2013.
- LeJOS Team (2008b) “NXJ technology”, <http://lejos.sourceforge.net/nxj.php>. Acessado em: 23/01/2013.
- LeJOS Team (2008c) “The LeJOS NXT tutorial”, Disponível em: <http://lejos.sourceforge.net/nxt/nxj/tutorial/index.htm>. Acessado em: 23/01/2013.
- Oliveira, Geraldo, Silva, Ricardo, Lira, Tiago, Reis, Luis Paulo. (2009) “Environment Mapping using the Lego Mindstorms NXT and leJOS NXJ”, Disponível em: <http://paginas.fe.up.pt/~ei04085/psite/LejosPaper.pdf>, Acessado em: 20/01/2013.
- Oliveira, Thiago Costa, Gonçalves, Nelson M. A., Ribeiro, João M. F. S. (2008) “Exploração da plataforma de programação leJOS para robôs Lego MindStorms: Uma Abordagem à Robótica Evolucionária”. Disponível em: http://www.di.uminho.pt/ensino/licenciaturas/lei/laboratorios-de-informatica-iv/47109_47119_47131.pdf. Acessado em: 18/01/2013.
- Hachour, O. (2008) “Path planning of Autonomous Mobile robo”, In: *International Journal Of Systems Applications, Engineering & Development*, 2ª Edição. <http://www.naun.org/multimedia/UPress/saed/saed-45.pdf>, Janeiro.

Preditor de números para lotofácil: uma abordagem usando algoritmos evolutivos

Fernando Barbosa¹, Lais C. R. S. Lopes¹, Mikaely Cruciol¹

¹Instituto Federal Goiano – Campus Urutaí – Departamento de Informática

Rodovia Geraldo Nascimento, Km 2,5 – Urutaí – GO – Brasil

{nando.cefet, laiscandidolopes, mikaelycruciol42}@gmail.com

Abstract. *Genetic algorithms (GA's) are the most widely used and studied evolutionary computing techniques. Based on the theory of evolution proposed by Darwin, has a powerful process of adaptation to certain types of problems, including global and local searches. This article aims, based on genetic algorithms, develop software predictor of numbers for a game of chance very popular in Brazil, Lotofácil. Maintained by the federal government this game reaches all sections of the population of the country.*

Resumo. *Os algoritmos genéticos (AG's) são as mais difundidas e estudadas técnicas de computação evolucionária. Baseado na teoria da evolução proposta por Darwin, possui um poderoso processo de adaptação a certos tipos de problema, inclusive buscas locais e globais. O presente artigo pretende, baseado em algoritmos genéticos, desenvolver um software preditor de números para um jogo de azar muito popular no Brasil, Lotofácil. Mantido pelo governo federal este jogo atinge a todas as camadas populacionais do país.*

1. Introdução

Os jogos de azar no Brasil são datados desde de 1784, o primeiro jogo que se tem registro teve o intuito de levantar verbas para a construção de obras públicas [Monteiro 2007]. Desde então essa modalidade de jogo caiu no gosto popular, além de representar uma forma de arrecadação para o governo federal.

A Caixa Econômica Federal é o órgão responsável por administrar as loterias federais no Brasil. Ela é mantenedora de nove jogos de azar, são eles, a Mega-Sena, Dupla-Sena, Quina, Lotéca, Lotogol, Lotomania, Timemania, Loteria federal e a Lotofácil. É com base no último jogo, Lotofácil, que o presente artigo propõe um preditor de números. A cartela da Lotofácil é composta por 25 números, onde o jogador pode marcar de 15 a 18 números, e ganhar caso acerte de 11 a 15 dezenas. O valor do prêmio máximo varia de acordo com o total arrecadado com a venda das cartelas, mais especificamente 46% do valor bruto arrecadado, sendo ainda deduzido deste os valores dos prêmios fixos, R\$2,50 para 11 acertos, R\$5,00 para 12 acertos e R\$12,50 para 13 acertos.

Todo tipo de jogo, seja ele de tabuleiro ou computacional, possui uma problemática, e uma forma de se resolver tais problemas é por meio de heurísticas. As heurísticas são métodos de investigação e solução de problemas baseados na busca algorítmica para

encontrar o ótimo ou uma solução que satisfaça determinados pré-requisitos. Elas procuram usar procedimentos, estratégias, aproximação, sempre na procura da melhor forma de se chegar a um determinado fim, garantindo assim soluções eficientes.

Uma heurística muito conhecida é o Algoritmo Genético (AG), inspirado na teoria de Charles Darwin, o qual afirma que mecanismos de evolução é uma competição que seleciona os indivíduos mais bem adaptados em seu ambiente, transmitindo características aos seus descendentes, permitindo sua sobrevivência. Os AGs são técnicas de busca para encontrar melhores soluções em problemas de otimização através da evolução de populações de soluções de cromossomos artificiais. Um cromossomo ou indivíduo é uma estrutura de dados que contém as possíveis soluções de determinado problema. Os cromossomos são submetidos a um processo evolucionário que envolve avaliação, seleção, cruzamento e mutação. Após vários ciclos de evolução, a população final é composta por indivíduos mais aptos [Pacheco 1999].

O presente artigo propõe um software preditor de números para Lotofácil. Vale ressaltar que é impossível prever com precisão quais números serão sorteados, pois, trata-se de um jogo com problemática de aleatoriedade. Quando a análise combinatória é empregada a esse problema tem-se que, a chance de alguém jogando 15 números acertar todos é de 3.268.760 para uma.

Este trabalho está estruturado da seguinte forma: metodologia, sessão em que será tratada a forma como o software foi confeccionado, bem como ferramentas utilizadas, nesta há ainda uma sub-sessão intitulada condução do experimento, que abordará a sistemática usada para realizar os testes. Resultados, sessão que demonstrará os resultados obtidos e por último tem-se a conclusão, sessão encerratória apresentado percepções finais acerca do software.

2. Metodologia

Todos os algoritmos e fórmulas, presentes no software, foram implementados por meio da plataforma @Java¹, usando-se o paradigma de orientação a objetos [Cardoso 2006].

Dentre a grande variedade de técnicas que anunciam sua eficiência em prever os números da Loteria Lotofácil, foi escolhida uma conhecida como *técnica das catorze dezenas*². Nesta as nove dezenas que mais saíram nos dez últimos jogos são selecionadas e reservadas. Posteriormente cinco dezenas são escolhidas, estas cinco não devem ter sido sorteadas no último jogo e nem estar entre o conjunto das nove dezenas, já separado. Juntando-se as nove mais saídas com as cinco escolhidas totaliza-se catorze dezenas distintas que serão fixas.

As onze dezenas restantes para completar os vinte e cinco números da cartela da Lotofácil, chamadas de dezenas variáveis, são concatenadas com as quatorze fixas. Dessa forma são gerados onze jogos diferentes, todos compostos pelas quatorze dezenas fixas e mais uma dezena das onze variáveis. É mostrado na Figura 1 um exemplo numérico de forma resumida.

Implementou-se a presente técnica para gerar onze jogos, e em seguida um algoritmo genético (AG) capaz de gerar uma população de jogos, cuja a distância fosse a

¹ Sítio oficial da linguagem java: <http://www.oracle.com/technetwork/java/index.html>

² Técnica das catorze dezenas, disponível no sítio: <http://www.loteriaspremium.com>

- Nove dezenas mais sorteadas:
1 – 8 – 12 – 15 – 20 – 22 – 23 – 24 – 25
- Cinco dezenas não sorteadas no último sorteio:
2 – 5 – 6 – 18 – 19
- Onze dezenas restantes:
3 – 4 – 7 – 9 – 10 – 11 – 13 – 14 – 16 – 17 – 21
- Jogos gerados:
1 8 12 15 20 22 23 24 25 2 5 6 18 19 3
1 8 12 15 20 22 23 24 25 2 5 6 18 19 4
...
1 8 12 15 20 22 23 24 25 2 5 6 18 19 21

Figura 1. Exemplo de implementação da técnica das 14 dezenas

menor possível da citada técnica.

Para o algoritmo genético (AG), foi gerada uma população inicial com quinhentos indivíduos, cada indivíduo ou cromossomo representa um jogo com quinze dezenas e mais o valor *fitness*, logo, uma matriz de quinhentas linhas por dezesseis colunas foi gerada para representar a população inicial. Cada cromossomo foi gerado com números aleatórios, respeitando o intervalo de 1 a 25 (números disponíveis na cartela da Lotofácil), tomou-se ainda o cuidado para que nenhum número do indivíduo fosse repetido. Na Figura 2 é demonstrada a representação de um indivíduo.



Figura 2. Representação de um cromossomo

A função de aptidão, considerada como o segundo passo de um AG, consiste em atribuir nota a um determinado cromossomo ou indivíduo. É com base nela que os cromossomos são selecionados e cruzados, a essa nota ou valor dá-se o nome de *fitness*. Como função de aptidão foi escolhido comparar as dezenas encontradas entre as duas técnicas. Para tanto, cada cromossomo foi comparado com os onze jogos gerados pela *técnica das quatorze dezenas*. Jogo a jogo as quinze dezenas entre os dois jogos foram confrontadas, toda vez que uma dezena presente no cromossomo estava também presente em um jogo gerado pela *técnica das quatorze dezenas* o valor do *fitness* aumentava em um ponto gerando um somatório. Quanto maior o *fitness* mais apto o indivíduo estava para 'adivinhar' os números da Lotofácil baseado na *técnica das quatorze dezenas*.

O cruzamento, terceiro passo do AG, é utilizado para recombinar material genético de dois pais previamente selecionados. Para selecioná-los foi utilizado o método de seleção por torneio [Lacerda and P. 2003], que consiste em atribuir um duelo entre dois ou mais pais prevalecendo o mais forte. O torneio foi aplicado entre dois pais escolhidos de forma aleatória, foram realizados duzentos e cinquenta duelos, metade da população inicial. Como de cada torneio um pai prevalecia, formou-se uma população de duzentos e cinquenta pais para participarem do cruzamento.

Para o cruzamento ou *crossover* os pares de pais selecionados foram escolhidos de forma aleatória. O operador de *crossover* aplicado foi o crossover de 1 ponto

[Lacerda and P. 2003], onde um ponto aleatório é fixado no filho sendo preenchido com genes de um pai até o ponto especificado e genes do outro pai após o ponto. De cada cruzamento foram gerados dois filhos. O processo de mutação, quarto passo do AG, serve para assegurar alguma diversidade entre os indivíduos. A mutação foi feita selecionando de forma aleatória dois genes em um determinado cromossomo e trocando seu material genético de posição, dessa forma testificava-se que não haveria inconsistência nos cromossomos, como números repetidos por exemplo. A fim de fazer um sistema mais dinâmico e flexível as taxas de mutação e quantidade de gerações ficaram a cargo do usuário, o que possibilita maior gama e formas de realização de testes, além de interação destes com o sistema.

2.1. Condução do experimento

Para gerar resultados legítimos, foi feito a descarga de todos os resultados existentes da Lotofácil, serviço oferecido pelo órgão mantenedor das loterias no Brasil, presente no sítio oficial do banco da Caixa Econômica ³.

Com esses resultados em mãos, foram selecionados, de forma aleatória, onze jogos sequenciais, dez serviram de entrada para o software preditor que gerou uma população ordenada de forma decrescente pelo seu *fitness*. O jogo que sobrou dos onze selecionados foi usado para verificar a quantidade de dezenas acertadas.

O indivíduo mais apto gerado pelo preditor foi comparado com o jogo reservado para realizar a verificação. A quantificação das dezenas acertadas foi feita de forma manual, onde cada dezena do indivíduo presente no jogo reservado para a verificação correspondia a um acerto. Todo o processo foi fielmente repetido quinze vezes, mudando apenas os dez jogos inseridos no sistema e o seu respectivo jogo usado como resposta. Na primeira execução dos testes, os jogos numerados de um a dez serviram de entrada para o sistema e o jogo de número onze foi usado com o intuito de realizar a verificação, dando sequência na segunda execução foram selecionados os jogos de número doze ao vinte e um, sendo o de número vinte e dois usado para a verificação e assim por diante.

Para a condução dos experimentos os parâmetros variáveis do AG foram setados de forma igual para todas as execuções, seus valores podem ser vistos na Tabela 1:

Tabela 1. Valores dos parâmetros utilizados para o preditor

Parâmetro	Valor
<i>pop_inicial</i>	500
<i>qtd_geracao</i>	200
<i>qtd_torneio</i>	250
<i>tx_mutacao</i>	5%

3. Resultados

Apesar de o foco do presente trabalho não tenha sido o de fazer um comparativo entre o preditor utilizando a *técnica das quatorze dezenas* e o preditor utilizando um AG, essa comparação foi feita, mas apenas com o intuito de enriquecer os resultados apresentados. Os resultados obtidos pela condução dos experimentos são mostrados na Tabela 2. Na

³Sítio oficial do banco Caixa: <http://www.caixa.gov.br/>

primeira coluna é exibido o número da execução, na segunda o número do concurso da Lotofácil utilizado para validar o desempenho, na terceira coluna é demonstrado a quantidade de dezenas acertadas pelo preditor utilizando um AG e na quarta e última coluna é exposto o desempenho do preditor aplicando a técnica das catorze dezenas somente, sem a utilização do AG. Nota-se que para gerar os dois resultados um mesmo jogo foi utilizado como validador.

Os resultados obtidos demonstraram que a técnica das catorze dezenas por si só obteve acertos inferiores se comparados aos acertos encontrados pelo AG. Embora o AG não tenha conseguido acertar os quinze números em nenhuma das quinze execuções, ele demonstrou resultados significativos sendo seu menor índice de acertos de cinquenta e três por cento e seu maior de noventa e três por cento. A média de dezenas acertadas pelo AG nas quinze execuções foi de 10,4 acertos por jogo o que equivale a 69,3%, enquanto que a média da técnica das 14 dezenas foi de 8,6 acertos por jogo, logo 57,3% .

Percebe-se que, o preditor quando usando a técnica das 14 dezenas não conseguiu em nenhuma das execuções acertar mais dezenas do que quando o preditor implementava o AG.

Tabela 2. Resultados obtidos nas execuções do preditor

execução	concurso	acertos - AG	acertos - s/ AG
1	11	11	10
2	22	10	08
3	33	09	09
4	44	10	08
5	55	09	09
6	66	10	10
7	77	10	06
8	88	14	08
9	99	10	10
10	110	08	08
11	121	13	12
12	132	11	10
13	143	10	10
14	154	11	07
15	165	10	06

4. Conclusão

O algoritmo genético se mostrou um preditor no mínimo interessante. Embora não tenha conseguido, durante a execução dos testes, prever as quinze dezenas de algum jogo obteve uma média de acertos consideravelmente boa, em torno de 10 dezenas por jogo.

É provável também que obtivesse resultados mais promissores se a quantidade de testes fosse aumentada, pois das quinze execuções realizadas, em uma o preditor utilizando o AG conseguiu prever 14 das 15 dezenas sorteadas. Vale lembrar que a probabilidade de se atingir 14 dezenas em um jogo da Lotofácil é de 21.791 para um, empregando o preditor esse número caiu para 8 tentativas. Seguindo esses resultados é de se esperar que,

ao realizar uma maior quantidade de teste, o preditor consiga acertar as quinze dezenas com uma probabilidade menor do que a existente, 3.268.760 para uma

Outro fator que pode ser levado em conta para melhorar ou piorar os resultados são as configurações dos parâmetros livres do algoritmo genético. Uma mudança na quantidade de gerações ou de indivíduos por população por exemplo poderia gerar resultados diferentes dos apresentados.

Como trabalhos futuros pretende-se desenvolver preditores aplicando outras técnicas bio inspiradas, ou mesmo, outro AG com técnicas de seleção, como o elitismo [Ricardo 2008] e o método da roleta [Ricardo 2008], para ao final fazer um comparativo e assim eleger o que melhor se adapta ao problema de prever os quinze números sorteados pela Lotofácil. Pretende-se ainda realizar estudos sobre análise combinatória e probabilidade para encontrar outras técnicas matemáticas válidas e assim gerar outras funções de aptidão.

Referências

- Cardoso, C. (2006). *Orientação a objetos na prática: aprendendo orientação a objetos com Java*. Ciência Moderna, 1th edition.
- Lacerda, E. G. M. and P., C. A. C. (2003). Introdução aos algoritmos genéticos.
- Monteiro, C. (2007). Howstuffworks - como funcionam as loterias. Disponível em <http://lazer.hsw.oul.com.br/loterias-brasil2.htm>. Acessado em 25 de janeiro de 2013.
- Pacheco, M. A. C. (1999). Algoritmos genéticos: Princípios e aplicações. ICA.
- Ricardo, L. (2008). *Algoritmos genéticos: uma importante ferramenta da inteligência computacional*. Brassport, 2th edition.

Interdisciplinaridades entre Pesquisa Operacional e Ciência da Computação

Bruno Felipe da Silva Rodrigues¹, João Paulo de Oliveira Santos¹, Stella Jacyszyn Bachega¹

¹Departamento de Engenharia de Produção – Universidade Federal de Goiás – Campus Catalão (UFG/CAC)

Av. Dr. Lamarthine Pinto de Avelar, 1120 – 75.704-020 – Catalão – GO – Brasil

brunofelipesr@hotmail.com, jotape_santos@hotmail.com.br, stella.bachega@catalao.ufg.br

Abstract. *Operations Research (OR) is an area of expertise that has large interdisciplinarity with various sciences. The aim of this paper is to demonstrate the interdisciplinarity among OR techniques and various segments of Computer Science. For this, we used the procedure of theoretical and conceptual research. The following OR techniques were analyzed: linear programming, nonlinear programming, dynamic programming, queuing theory, simulation, heuristics and stochastic methods. We verified applications of these techniques in computer networks, cryptography, computer vision, artificial intelligence, information systems, operating systems, algorithm development and software engineering.*

Resumo. *A Pesquisa Operacional (PO) é uma área do conhecimento que possui grande interdisciplinaridade com diversas ciências. Esse artigo possui o objetivo de evidenciar interdisciplinaridades das técnicas de PO em trabalhos desenvolvidos em diversos segmentos da Ciência da Computação. Para tanto, utilizou-se o procedimento de pesquisa teórico-conceitual. As técnicas de PO analisadas foram: programação linear, programação não linear, programação dinâmica, teoria das filas, simulação, heurísticas e métodos estocásticos. Verificaram-se aplicações dessas técnicas em redes de computadores, criptografia, visão computacional, inteligência artificial, sistemas de informação, sistemas operacionais, desenvolvimento de algoritmos e engenharia de software.*

1. Introdução

A Pesquisa Operacional (PO) é um termo que surgiu durante a Segunda Guerra Mundial, utilizado para designar o desenvolvimento de métodos científicos relacionados à utilização mais eficaz dos recursos e à resolução de problemas militares, como alocação de navios, tipo de avião a utilizar em determinado ataque, dimensionamento de comboios, entre outros (ARENALES et al, 2007).

Após o notável sucesso das aplicações de PO na resolução de problemas militares, diversas áreas do conhecimento iniciaram a sua inserção nos ambientes acadêmicos e empresariais (ANDRADE, 2004). A partir do desenvolvimento de estudos e aplicações de métodos nesses novos cenários, começaram a surgir conceitos de diversas técnicas dessa área, como programação linear, programação não linear, teoria das filas, simulação computacional, entre outros, que viriam a ser técnicas importantes e

utilizadas na tomada de decisões para a resolução de problemas. Dentre as áreas de aplicação da PO está a Ciência da Computação.

Visto a enorme abrangência da Pesquisa Operacional, esta pesquisa se faz necessária para esclarecer não apenas aos estudantes de graduação, como a toda a comunidade acadêmica, sobre as principais definições e características da área e suas possíveis interdisciplinaridades. Portanto, o objetivo desse artigo é evidenciar interdisciplinaridades das técnicas de PO em trabalhos desenvolvidos em diversos segmentos da Ciência da Computação, apontando assim possíveis vínculos entre essas áreas.

Para cumprir esse objetivo, o presente artigo é estruturado da seguinte forma: na segunda seção são apresentadas brevemente as principais técnicas de pesquisa operacional; na terceira seção há a metodologia de pesquisa; na quarta seção têm-se as interdisciplinaridades observadas entre PO e Ciências da Computação; e na quinta seção há as considerações finais.

2. Técnicas de Pesquisa Operacional

A seguir são expostas brevemente as principais técnicas de PO, a saber:

Programação linear

A Programação Linear (PL) soluciona problemas que dizem respeito, principalmente, a alocação ótima de recursos escassos para realização de atividades (COLIN, 2007).

Os estudos de programação linear se utilizam de modelos matemáticos na resolução de problemas. Os modelos de otimização linear têm a característica de que todas as funções-objetivo e as restrições impostas pelos problemas são representadas por funções lineares. Entretanto, essas não são as únicas particularidades da técnica. Conforme Colin (2007, p. 07), há quatro características fundamentais para os problemas de programação linear: proporcionalidade, aditividade, divisibilidade e certeza.

Programação não linear

De acordo com Lachtermacher (2004, p. 292), a Programação Não Linear (PNL) trata de “problemas de otimização em que a função-objetivo e/ou pelo menos uma das restrições envolvidas não são funções lineares das variáveis de decisão”.

Assim, a PNL é mais complexa do que a PL devido a não-linearidade que engloba um universo infinito de funções de distintas formas. Hillier e Lieberman (2006) exemplificam alguns tipos de problemas de PNL, a saber: otimização irrestrita, otimização linearmente restrita, programação quadrática, programação convexa, programação separável, programação não-convexa, programação geométrica, programação fracionária e problema da complementaridade.

Programação dinâmica

A Programação Dinâmica (PD) é uma técnica de programação considerada extremamente útil e versátil, pois pode ser utilizada na resolução de diversos tipos de problemas, tais como: lineares ou não lineares, determinísticos ou estocásticos, com horizonte finito ou infinito (COLIN, 2007).

Segundo Hillier e Lieberman (2006, p. 424), a PD “[...] é uma técnica matemática útil para criar uma sequência de decisões inter-relacionadas”. Portanto, pode ser considerada uma técnica de decomposição que utiliza recursividade na resolução dos problemas, onde o problema principal é dividido em subproblemas. Os problemas menores são otimizados e suas resoluções são guardadas para, posteriormente, serem combinadas na resolução do problema principal. Tal abordagem pode ser descrita como uma técnica de *bottom-up*.

Teoria das filas

Arenales *et al.* (2007, p. 433) define a Teoria das Filas como: “um ramo da Pesquisa Operacional que estuda as relações entre as demandas em um sistema e os atrasos sofridos pelos usuários deste sistema”. O autor enfatiza que as filas são formadas quando a capacidade do sistema de fornecer o serviço, dentro de um determinado tempo, é menor do que a demanda.

A análise de sistemas de filas busca quantificar o fenômeno de espera em filas, ou seja, determinam-se medidas de desempenho da espera em filas, como tempo médio de espera para receber um serviço, produtividade da instalação de um serviço, ou o tamanho médio de uma fila (TAHA, 2010).

Simulação

Simulação é um processo de experimentação com um modelo detalhado de um sistema real para determinar como o sistema responderá a mudanças em sua estrutura, ambiente ou condições de contorno (HARREL *et al.*, 2002). Os diversos usos dessa técnica podem ser agrupados em três grandes categorias (BUFFA; SARIN, 1987): i) projeto, ii) diagnóstico, e iii) treinamento.

A simulação pode ser realizada com auxílio de computadores. Conforme Law e Kelton (2000), a simulação computacional é constituída por técnicas que usam computadores para “imitar” ou simular diversos tipos de operações ou processos do mundo real.

Heurísticas

Heurísticas são conjuntos de técnicas e algoritmos computacionais para a resolução de problemas que, segundo Hillier e Lieberman (2006, p. 599), provavelmente encontrarão soluções viáveis, mas não necessariamente uma solução ótima, para dado problema.

As Heurísticas surgiram em resposta à natureza complexa dos problemas de otimização com programas lineares, não lineares e inteiros. Utilizando esforço computacional, as heurísticas garantem a viabilidade ou a otimalidade da solução encontrada (REHFELDT, 2001).

Processos estocásticos

De acordo com Hillier e Lieberman (2006, p. 713), “um processo estocástico é uma coleção de variáveis aleatórias indexadas (X_t), onde t é um índice definido num conjunto T .” Assim, um processo estocástico é a descrição de um fenômeno aleatório que varia com o tempo. Possuem importância em diversas áreas do conhecimento, pois descrevem o comportamento de um sistema que opera ao longo de algum período.

Arenales *et al.* (2007, p. 407) advogam que os processos estocásticos têm sido aplicados em diversas áreas onde alguns dados não são conhecidos com certeza, como em confiabilidade de sistemas, manutenção de máquinas etc.

3. Metodologia da Pesquisa

A pesquisa teórico-conceitual pode ser considerada como uma pesquisa bibliográfica que, na visão de Silva e Silva (2004), possui o intento de explicar um problema a partir de referências teóricas publicadas em documentos, podendo ser efetuada como parte da pesquisa descritiva ou experimental ou, ainda, de forma independente.

De acordo com Berto e Nakano (1998; 2000), esse tipo de pesquisa é fruto de uma série de reflexões fundamentadas em um fato observado ou exposto pela literatura, reunião de opiniões e ideias de diversos autores ou mesmo pela simulação e modelagem teórica. Conforme esses autores, as discussões conceituais baseadas na literatura e revisões bibliográficas são classificadas como pesquisas teórico-conceituais. Neste trabalho foi realizada a pesquisa teórico-conceitual, ou bibliográfica, com o intuito de pré-orientação teórica.

Para tal revisão, foram consultadas bases de dados, sites de busca, bancos de teses e dissertações de universidades nacionais, e anais dos eventos. Dentre eles estão: Periódicos CAPES, SciELO, Google Scholar, Bancos de teses da USP, UFSCar, UNICAMP e UFSC, Anais do ENEGEP, SBPO e SEGET. Dentre as palavras-chave utilizadas estão os nomes completos das técnicas de pesquisa operacional, suas siglas, áreas de ciência da computação, e suas combinações.

Ressalta-se que, no presente momento, a pesquisa bibliográfica não é considerada exaustiva e que será ampliada no decorrer da pesquisa. Portanto, os resultados aqui apresentados mostram algumas das possíveis aplicações de técnicas de PO em algumas áreas da Ciência da Computação. Ainda, as técnicas de PO apresentadas nesse artigo foram selecionadas por serem técnicas mais divulgadas nas referências base da área.

4. Interdisciplinaridades Observadas

Nessa seção há a apresentação de interdisciplinaridades notadas entre técnicas de pesquisa operacional e áreas da ciência da computação.

4.1. Programação linear

A PL tem grande interdisciplinaridade com várias áreas do conhecimento, ela pode ser usada na computação para determinar, por exemplo, o número máximo de processos que uma máquina pode realizar, submetido a restrições de capacidades e de tempo (COLIN, 2007).

Longo (2004) explorou técnicas mais recentes de programação linear inteira em aplicações de problemas de roteamento de veículos, preocupando-se com algoritmos que tenham tempos computacionais mais aceitáveis para encontrar a solução ótima para instâncias de porte razoável.

Algoritmos de PL que preservem a privacidade para utilização em computação segura multiparte foi o foco do trabalho de Deitos (2009), o qual aplicou em problemas de otimização da cadeia logística.

Mapa e Lima (2012) usaram de forma combinada sistemas de informações geográficas para transporte (SIG-T) com programação linear inteira mista na resolução de problemas de localização de instalações.

Portanto, puderam-se notar aplicações de PL nas áreas de criptografia computacional e sistemas de informação.

4.2. Programação não linear

Dentre os usos de PNL está o de Facó e Oliveira (2002; 2003). Esses autores aplicaram computação paralela para resolução numérica de problemas de PNL com restrições. Para realização do trabalho, utilizaram o método do Gradiente Reduzido Generalizado (GRG) e o código computacional LSGRG2.

Vasconcelos (2007) desenvolveu uma metodologia para avaliação e controle de segurança de sistemas elétricos interligados com base em métodos de aprendizagem automática. No decorrer da pesquisa foi utilizado um problema de otimização que inclui ANN (*Artificial Neural Networks* – Redes Neurais Artificiais) como restrição funcional, o que caracterizando um modelo de PNL.

Assim, nessa subseção foram observados usos de PNL em redes de computadores e inteligência artificial.

4.3. Programação dinâmica

Uma aplicação direta da PD na computação é o desenvolvimento de algoritmos *cache-oblivious* para problemas que utilizam programação dinâmica. Como explicitado em Rodrigues (2008), a PD divide seus problemas em problemas menores e necessita de boa memória para guardar as soluções desses subproblemas, o algoritmo *cache-oblivious* se mostrou eficiente no uso da memória *cache*.

Outro trabalho foi a análise e implementação do algoritmo de PD de Nussinov-Jacobson para otimização das diversas características da estrutura secundária do RNA (ou Ácido Ribonucleico), descrito em Rockenbach (2009). Logo, foram identificadas aplicações de PD nas áreas de sistemas operacionais e desenvolvimento de algoritmos.

4.4. Simulação

A simulação é uma ferramenta eficaz e flexível para a modelagem de sistemas. Uma aplicação dessa técnica de PO foi feita por Carvalho (2005) para estudo de melhoria do desempenho de redes de computadores.

Oliveira e Silva (2007) desenvolveram uma pesquisa para construção de um sistema que utiliza simulação e interface gráfica para representar o conceito de máquina de Turing. A interface, construída em Java, foi baseada em técnicas de programação de jogos e possuiu a finalidade principal de auxiliar o processo de ensino-aprendizagem em disciplinas como Teoria da Computação.

Portanto, foram identificadas aplicações de simulação em redes de computadores e desenvolvimento de algoritmos.

4.5. Teoria das filas

Na Ciência da Computação a maioria das aplicações dessa ferramenta da Pesquisa Operacional são processos de algoritmos computacionais esperando para utilizar uma CPU ou um servidor, formando filas que podem ser estudadas pela teoria.

A teoria das filas pode ser utilizada para analisar o desempenho de sistemas de atendimento. Carrión (2007) desenvolveu um estudo em um servidor utilizando essa técnica de PO. Em tal pesquisa, houve a otimização do congestionamento no servidor causado pela enorme quantidade de mensagens que chegam ao mesmo tempo. Entretanto, devido à complexidade em descrever o comportamento do tráfego digital, caracterizado pela distribuição exponencial, diferentes modelos de filas e métodos matemáticos avançados, como distribuições em cauda longa e a transformada de Laplace, foram desenvolvidos para tentar solucionar o problema. Assim sendo, identificou-se um uso de teoria das filas em redes de computadores.

4.6. Heurísticas

As heurísticas são utilizadas quando o tempo requisitado para encontrar a solução de um problema é menor se comparado ao enorme tempo gasto por algumas das técnicas de solução, como a programação dinâmica e os algoritmos *branch-and-bound* e *branch-and-cut*. Assim, os esforços na área de Ciências de Computação estão na elaboração de algoritmos que resolvam, cada vez melhor, problemas complexos, com uso dessa técnica de PO. Exemplos de métodos heurísticos podem ser encontrados em Prestes (2006) para solucionar o Problema do Caixeiro Viajante – PCV. No estudo, diversas abordagens heurísticas, como a Busca Tabu, o Simulated Annealing (SA) e o Greedy Randomized Adaptive Search Procedures (GRASP), são utilizadas para criar algoritmos mais eficientes para solucionar o PCV.

Outros exemplos de aplicações de algoritmos desenvolvidos por meio de heurísticas também podem ser encontrados em: Rehfeldt (2001) e Toffolo *et al.* (2006) utilizaram heurísticas para desenvolverem algoritmos que otimizassem o escalonamento de tarefas, o primeiro de tripulações aéreas enquanto que o segundo na indústria de calçados; e Pileggi *et al.* (2007) utilizaram heurísticas para gerar um algoritmo que abordasse o sequenciamento de padrões de cortes bidimensionais na indústria automobilística.

Logo, foram verificados usos de heurísticas nas áreas de sistemas de informação e desenvolvimento de algoritmos.

4.7. Métodos estocásticos

Domenech e De Biasi (2012) descreveram a arquitetura geral de um sistema e as características principais de um protótipo de ferramenta de software destinada a auxiliar estudantes durante o processo de ensino-aprendizagem de determinados métodos estocásticos como Cadeias de Markov.

Fernandes (2010) aplicou a Estereoscopia (visão estéreo), que é um processo em Visão Computacional 3-D juntamente com métodos estocásticos. Foi utilizada a abordagem estocástica para o processo de combinação de imagens, detalhando três algoritmos de otimização aplicada à correspondência estérea.

Assim, notou-se aplicação de métodos estocásticos em engenharia de software e visão computacional.

5. Considerações finais

O objetivo almejado nesse artigo foi alcançado. Por meio de uma pesquisa teórico-conceitual, evidenciaram-se possíveis interdisciplinaridades entre técnicas de pesquisa operacional com aplicações em diversos segmentos da Ciência da Computação.

Dentre os vínculos das técnicas de PO analisadas com áreas da ciência da computação estão: redes de computadores (aplicações de programação não linear, simulação e teoria das filas), criptografia computacional (aplicação de programação linear), engenharia de software (aplicação de métodos estocásticos), visão computacional (aplicação de métodos estocásticos), inteligência artificial (aplicação de programação não linear), sistemas de informação (aplicações de programação linear e heurísticas), sistemas operacionais (aplicação de programação dinâmica), desenvolvimento de algoritmos (aplicação de programação dinâmica, simulação e heurísticas).

O presente trabalho contribui ao propiciar maior oportunidade para que estudantes de ciência da computação possam focar sua formação acadêmica nas bases da Pesquisa Operacional, desenvolvendo, assim, novas pesquisas e projetos interdisciplinares. Salienta-se que o presente trabalho faz parte de uma pesquisa que visa aplicar técnicas de pesquisa operacional em determinadas áreas da Ciência da Computação.

Referências

- Andrade, E. L. (2004) “Introdução à Pesquisa Operacional: Métodos e Modelos para Análise de Decisões”, LTC, 3ª edição.
- Arenales, M.; Armentano, V.; Morabito, R.; Yanasse, H. H. (2007) “Pesquisa Operacional para cursos de engenharia”, Campos.
- Berto, R. M. V. S.; Nakano, D. N. (1998). Metodologia da pesquisa e a engenharia de produção. In: *XVIII Encontro Nacional de Engenharia de Produção (ENEGEP) e IV International Congress of Industrial Engineering (ICIE)*, 1998, UFF/ABEPRO, out.
- Berto, R. M. V. S.; Nakano, D. N. (2000). A produção científica nos anais do encontro nacional de engenharia de produção: um levantamento dos métodos e tipos de pesquisa. In: *Produção*, v. 9, nº 2, p. 65-75, jul.
- Bornatto, G. (2002). Modelagem - Simulação - Informática e a Matemática. In: *Rev. PEC*, Curitiba, v.2, n.1, p.67-71, jul. 2001-jul. 2002
- Buffa, E. S.; Sarin, R. K. (1987) “Modern Production/Operations Management”, John Wiley & Sons, 8ª ed.
- Carrión, E. A. (2007) “Teoria das Filas Como Ferramenta para Análise de Desempenho de Sistemas de Atendimento: Estudo de Caso de um Servidor da UECE”, CCT/UECE/CEFET.
- Carvalho, E. M. (2005). “Modelagem e Simulação de Desempenho de Redes de Computadores”, DC/UEL.
- Colin, E. C. (2007) “Pesquisa Operacional: 170 aplicações em Estratégia, Finanças, Logística, Produção, Marketing e Vendas”, LTC.
- Deitos, R. J. (2009). “Algoritmos de Programação Linear com Atributos de Privacidade para o uso em Computação Segura Multi-Parte”, UFSC.
- Domenech, M.; De Biasi, H. (2012). Ferramenta de Software para o Auxílio ao Processo de Ensino-Aprendizagem de Métodos Estocásticos. In: *UNOESC & Ciência*, v. 3, nº 1, p. 37-46, jan./jun.

- Faco, J. L. D.; Oliveira, L. A. A. (2003). Aplicação de Computação paralela à resolução numérica de problemas de programação não linear com restrições. In: XXXV *Simpósio Brasileiro de Pesquisa Operacional*, 2003, SBPO.
- Fernandes, J. L. (2010). Aplicação de Algoritmos de Otimização Estocástica à Correspondência entre Imagens em Visão Estéreo. In: 9th Brazilian Conference on Dynamics, Control and their Applications June 07-11, 2010.
- Harrel, C. R.; Mott, J. R. A.; Bateman, R. E.; Bowden, R. G.; Gogg, T. J. (2002) “Simulação: otimizando os sistemas”, Belge Simulação e IMAM, 2^a ed.
- Hillier, F. S.; Lieberman, G. J. (2006) “Introdução à Pesquisa Operacional”, McGraw Hill, 8^a edição.
- Law, A. M.; Kelton, W. D. (2000) “Simulation Modeling & Analysis. McGraw- Hill, 3rd edition.
- Longo, H. J. (2004) “Técnicas para programação inteira e aplicações em problemas de roteamento de veículos”, PUC-RJ.
- Mapa, S. M. S.; Lima, R. S. (2012). Uso combinado de sistemas de informações geográficas para transportes e programação linear inteira mista em problemas de localização de instalações. In: *Gest. Prod.*, v.19, n.1, p. 119-136.
- Morais, M. L.; Santos, S. A. (2011). Análise Comparativa para o Problema de Localização-Alocação: Modelo Não Linear Geral Versus Modelo das P-Medianas com Variáveis Inteiras: Um Estudo de Caso, IMECC/UNICAMP.
- Oliveira, G. P.; Silva, M. A. S. (2007). Construção de Simuladores Gráficos para Teoria da Computação: Uma Proposta para o Ensino do Conceito de Máquinas de Turing. In: IV Simpósio de Excelência em Gestão e Tecnologia (SEGET), 2007, AEDB.
- Pileggi, G. C. F.; Morabito, R.; Arenales, M. N. (2007). Heurísticas Para os Problemas de Geração e Sequenciamento de Padrões de Corte Bidimensionais. In: *Pesquisa Operacional*, v. 27, nº 3, p. 549-568, set.
- Prestes, A. N. (2006) “Uma Análise Experimental de Abordagens Heurísticas Aplicadas ao Problema do Caixeiro Viajante”, PPGSC/UFRGN.
- Rehfeldt, M. J. H. (2001) “Uma Heurística Aplicada a um Problema de Escalonamento na Indústria Calçadista”, PPGA/UFRGS.
- Rockenbach, R. (2009) “Programação Dinâmica Aplicada a Problemas de RNA-Folding”, CCET/UNIOESTE
- Rodrigues, F. C. (2008) “Programação Dinâmica Eficiente com Algoritmos Cache-Oblivious”, IICC/UFRGS.
- Silva, H. H. R.; Silva, M. L. A. (2004) “Metodologia da pesquisa”, Salesiano.
- Taha, H. A. (2010) “Pesquisa Operacional”, Pearson, 8^a edição.
- Tofollo, T. A. M.; Souza, M. J. F.; Pontes, R. C. V.; Silva, G. P. (2006). Heurística de Recobrimento Aplicada à Escala de Tripulações Aéreas. In: XXXVIII *Simpósio Brasileiro de Pesquisa Operacional (SBPO)*, 2006, UCG/SOBRAPO, set.
- Vasconcelos, M. H. O. P. (2007) “Avaliação e Controlo de Segurança de Redes Interligadas com Grande Penetração Eólica com base em Métodos de Aprendizagem Automática”, Universidade do Porto/FEUP.

OpenNI e Suas Aplicações

Lorena A. Rezende, Dalton M. Tavares

Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Avenida Dr. Lamartine Pinto de Avelar, 1120, Setor Universitário, Catalão–GO – Brasil

{lorennarezende}@live.com, {dalton.tavares}@catalao.ufg.br

Abstract. *This paper aims to show the results of the combination between the framework OpenNI and the Kinect sensor, used to assist the development of applications based on natural interaction. The use of controllers like keyboards, mouse and remote controls to manage an application is no longer interesting and therefore, are getting obsolete. This opens the way to a more natural interaction between human and machines.*

Resumo. *Este artigo tem como objetivo apresentar o resultado da combinação do framework OpenNI com o sensor kinect, usado para auxiliar no desenvolvimento de aplicações baseadas em interação natural. O uso de controladores como teclados, mouses e controles remotos para manipular uma aplicação já não é mais tão interessante e por isso, estão se tornando obsoletos. Isso abre o espaço a uma interação mais natural entre humano e máquinas.*

1. Introdução

O OpenNI¹ (*Open Natural Interface*) é uma organização sem fins lucrativos que disponibiliza um SDK (*Software Development Kit*) de código aberto (*open source*), o que o torna um *framework* muito útil no auxílio ao desenvolvimento de interfaces (entre dispositivos, aplicações e bibliotecas de *middleware*) para aplicações baseadas em interação natural (*natural interface* ou NI), principalmente por meio da audição e visão.

O OpenNI suporta aplicações em 3 dimensões. Um exemplo da sua utilização é no auxílio à engenheiros de turbina, para otimizar seus projetos de forma interativa, permitindo que os usuários possam modificar modelos, re-executar simulações e inspecionar os resultados. Nesse contexto, a plataforma OpenNI e o *middleware* NITE são responsáveis por processarem a profundidade de imagens. A aplicação reconhece as movimentações da mão para modificar lâminas de turbina, em parâmetros como “direita/esquerda” para girar, “cima/baixo” para rotação/elevação, “frente/trás” para aumentar ou diminuir *zoom* e, quando a mão é aberta, a aplicação reconhece que é para parar o processo de interpretação [Rogge et al. 2012].

O Kinect é um sensor desenvolvido em 2010 pela Microsoft® para o console de jogos Xbox 360, o qual permite ao usuário interagir de diversas maneiras (ex. por meio de gestos, movimentos ou voz) com um jogo multimídia. Por possuir sensores de profundidade e também sensores que permitem executar o rastreamento de movimentos em tempo real, esse dispositivo vem ganhando destaque junto à comunidade científica com aplicações nos campos de robótica, interface humano computador e computação

¹Maiores informações em: <http://www.openni.org/about> Acesso em: 23/01/2013.

gráfica. Um exemplo da utilização do Kinect é o desenvolvimento de uma ferramenta chamada “*Hand-Potter*” [Moore 2012], muito útil no contexto da criação de objetos virtuais. A *Hand-Potter* utiliza o Kinect e algoritmos avançados que permitem a uma pessoa criar objetos virtuais em 3 dimensões, usando apenas suas mãos. Os algoritmos reconhecem as mãos e identificam sua interatividade. Dessa maneira, ele projeta virtualmente, exatamente a forma que foi “desenhada”.

O Kinect foi escolhido como dispositivo de interface com o intuito de explorar os recursos oferecidos pelo OpenNI, devido a capacidade inata desse dispositivo quanto ao reconhecimento de noção de profundidade, gestos, imagens, áudio, além de dar a robôs/aplicações a visão em 3 dimensões. No decorrer desse artigo serão apresentados os requisitos para a utilização do *framework* OpenNI e uma demonstração de sua utilização usando um sensor Kinect.

Esse artigo está organizado da seguinte forma: a seção 2 irá discutir o *framework* OpenNI, a seção 3 apresenta o sensor Kinect, a seção 4 mostra um breve estudo de caso e a seção 5 são apresentadas a conclusão e propostas de trabalhos futuros.

2. O *Framework* OpenNI

O OpenNI SDK é um kit de desenvolvimento de software *open source*, suportado em múltiplas plataformas. Além deste, existe também o OpenNI *framework*², utilizado como auxílio ao desenvolvimento de aplicações baseadas em interação natural.

O OpenNI possui uma variedade de aplicações pré-compiladas, que podem ser executadas logo após a instalação. Seu objetivo é padronizar a compatibilidade e interoperabilidade da interação natural, dispositivos, aplicações e *middlewares*. É compatível com suas versões anteriores, de modo que, uma aplicação feita em uma versão anterior pode ser executada usando a versão mais nova do OpenNI, sem precisar ser recompilada.

Também permite que aplicações sejam escritas sem que os desenvolvedores se preocupem com o fornecedor da tecnologia que habilita uma determinada característica (ex. visão 3D). Define os nós de produção, podendo ser do tipo *Device* (sensor), *Depth* (gera mapas de profundidade), *Image* (gera mapas de imagem-cor), *IR* (gera mapas de imagem infra-vermelho), *Audio* (gera um fluxo de áudio), *Gestures* (gera retornos quando gestos específicos são identificados), *SceneAnalyser* (analisa o cenário, separando primeiro plano/fundo), *Hand* (gera retorno quando pontos de mão são identificados) e *User* (gera a representação de um usuário no espaço 3D). O *framework* é responsável por traduzir dados brutos em dados “de alto nível”, transformando-os em informações que sejam mais fáceis de serem compreendidas por uma biblioteca de *middleware*. No contexto deste, ele será usado para auxiliar a execução de uma aplicação já pré-compilada.

O *framework* é baseado em eventos, realizando *callbacks* sempre que um evento ocorre. Permite que dados da execução de uma aplicação sejam gravados para serem reproduzidos posteriormente, funcionando como uma simulação. No contexto deste, se destina na interface dispositivo-*middleware*/aplicação suportando dados 3D.

²Para maiores informações consulte a documentação OpenNI disponível em:
<http://www.openni.org/openni-sdk/openni-sdk-history-2> Acesso em: 23/01/2013.

3. Kinect

O Kinect é um sensor de 3 dimensões (3D) desenvolvido em 2010 com a ajuda de Andrew Black (IEEE) e sua equipe da Microsoft Research Cambridge, bem como a empresa de semicondutores PrimeSense (tecnologia que possibilita aos dispositivos observar um cenário em 3 dimensões) [Ackerman 2011]. O Kinect possui sensores RGB, infra-vermelho, CMOS e dois microfones embutidos. Foi projetado para o console XBOX 360, com o objetivo de que o controlador seja o próprio jogador (a). Dessa forma, o jogador (a) é responsável por controlar a execução de um game por meio de movimentos corporais. O Kinect consegue abstrair as informações do ambiente usando o laser do sensor infra-vermelho para determinar a distância para cada pixel. Essa informação é mapeada sobre uma imagem a partir de uma câmera RGB padrão. Cada pixel tem uma cor e uma distância, podendo então, ser utilizado para mapear posições do corpo, gestos, movimentos ou ainda gerar mapas 3D [Wang et al. 2012].

A Microsoft lançou o Kinect para o console Xbox 360, porém, por ser robusto e de baixo custo, vem sendo utilizado em pesquisas nos campos de robótica e interface humano-computador. No contexto deste, o uso do Kinect se destina a reconhecer um usuário em cena e rastrear seus movimentos.

4. Estudo de Caso

Para utilizar os recursos oferecidos pelo *framework* OpenNI, primeiramente, é necessário ter instalado os seguintes componentes³:

- GCC 4.x;
- Python 2.6+/3.x;
- LibUSB 1.0.x;
- FreeGLUT3;
- JDK 6.0;
- OpenNI 1.5.x.x.

Os componentes supracitados são requisitos para se usar os recursos do *framework* OpenNI. Além disso, é necessário instalar o *middleware* NITE e o Sensor PrimeSense de modo a habilitar a operação do Kinect. O NITE é uma biblioteca de *middleware* disponibilizada pela PrimeSense que contém algoritmos para rastreamento de movimentos. Utilizando um sensor, como o Kinect, torna-se possível rastrear os movimentos de um ponto da mão (*hand-point*) ou ainda, movimentos do corpo (*full-body*). O sensor PrimeSense é uma abstração necessária para a aquisição de dados de um dispositivo e calcular a profundidade.

Quando uma aplicação é executada, os dados são gerados por um dispositivo físico que possui os sensores necessários para sua execução. Como a aplicação criada (seção 4.1 e 4.2) é baseada em 3D, a ferramenta utilizada foi o Kinect. Em nível de hardware, ele possui os sensores necessários: RGB, infra-vermelho (IR) e CMOS.

³ Disponível em: <http://www.openni.org/Downloads/OpenNIModules.aspx>. Acesso em: 10/01/2013.

Existe uma camada (SoC – *System on Chip*)⁴ que opera sobre o hardware descrito. Esta é responsável por calcular a profundidade, onde um sensor infra-vermelho ilumina a cena e um sensor CMOS decifra a luz recebida pelo infra-vermelho e produz uma imagem de profundidade VGA.

O *framework* OpenNI permite que o *host* (máquina) opere os sensores e acesse todos os dados brutos gerados por ele. O *framework* então, os traduz para informação de alto nível e os envia para a biblioteca de *middleware* NITE, a qual possui os algoritmos de rastreamento necessários para a execução da aplicação. Uma representação da arquitetura descrita é apresentada na Figura 1.

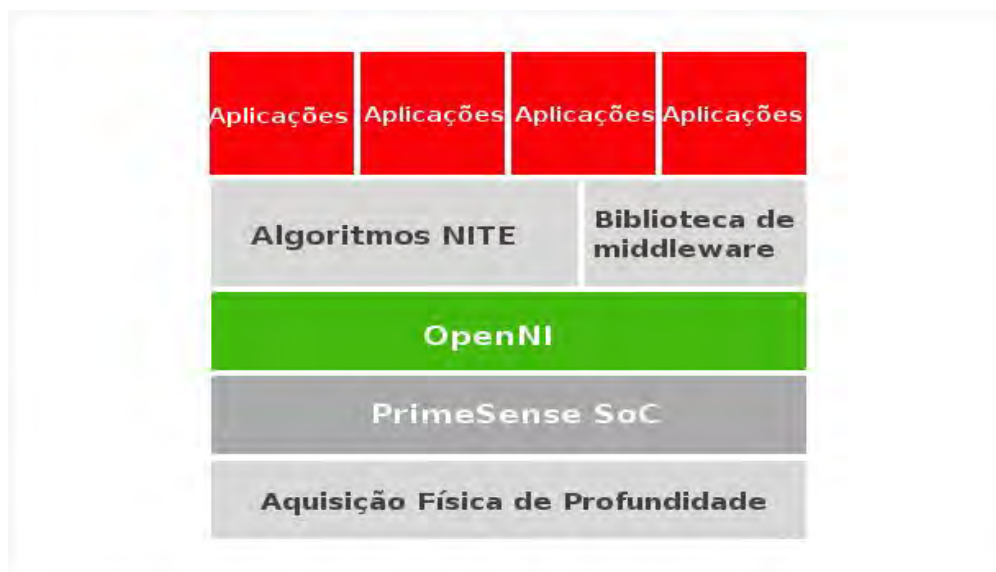


Figura 1. Modelo de representação de uma aplicação com interface natural baseada em dados 3D.⁵

4.1. Aplicação Desenvolvida

A aplicação criada é responsável por identificar um usuário em cena e rastrear seus movimentos. Para sua execução é utilizado o *framework* OpenNI, a biblioteca de *middleware* NITE e como dispositivo físico, o sensor Kinect. A aplicação possui três porções principais representadas pelos seguintes arquivos⁶: uma biblioteca “*opengles*” que permite a reprodução gráfica dos resultados obtidos, o “*SceneDrawer*” que processa os eventos (desenha) e o programa principal “*main*”, onde são realizadas as chamadas de funções e *callbacks*.

No programa principal (*main*), são declaradas variáveis globais dos tipos *Context*, *ScriptNode*, *DepthGenerator*, *UserGenerator* e *Player*. O *Context* é um objeto usado para criar os gráficos de produção. Estes são representados por um conjunto de nós que foram produzidos, bem como outros nós os quais são dependentes. Por exemplo, para obter um nó do tipo *UserGenerator* é necessário que exista um nó

⁴Mais informações em: <http://www.primesense.com/solutions/technology> Acesso em: 28/01/2013.

⁵Imagem original disponível em: <http://www.primesense.com/solutions/nite-middleware> Acesso: 29/01/2013.

⁶Para maiores informações consulte a documentação OpenNI Compliant disponível em: <http://www.openni.org/openni-sdk/openni-sdk-history-2> Acesso em: 23/01/2013.

DepthGenerator. No caso de haver mais de um sensor com a mesma funcionalidade, a aplicação pode escolher o gráfico de produção preferido (estes dispositivos podem gerar dados diferentes) levando em consideração um fornecedor específico, por exemplo. O *ScriptNode* é um objeto carregado a partir de uma *string* ou arquivo de script XML, usado para a execução do gráfico de produção. O *DepthGenerator* gera um mapa de profundidade, onde cada pixel no mapa tem um valor que representa a distância do mesmo para o sensor. O *UserGenerator* é um nó gerador de dados referentes à informações dos usuários em um cenário, onde cada usuário é identificado como único. Por fim, o nó *Player*, armazena a gravação dos dados de uma sessão que gera dados.

O OpenNI usa contagem por referência para que os objetos sejam compartilhados entre vários componentes da aplicação. Quando um tipo específico de nó é criado, sua referência é setada com o valor 1 e o identificador é retornado ao criador. A contagem de referência é incrementada (+1) no caso de haver novos nós que possuem o mesmo tipo de um nó já existente. Quando os nós não são mais necessários, vão sendo removidas/decrementadas (-1) as referências (por meio da função *void CleanupExit()*) até que ela tenha valor 0 e então, o nó é destruído e a memória antes alocada por ele, é liberada. Como o OpenNI suporta gravações de uma sessão geradora de nós, ao executar a aplicação, se não é passado nenhum parâmetro (o nome de um arquivo de gravação), o gráfico de produção será criado de acordo com um arquivo XML padrão do OpenNI. Caso contrário, ele iniciará a partir do arquivo de gravação especificado.

No caso apresentado, não foi passado nenhum parâmetro. Dessa forma, a execução do programa inicia normalmente, esperando por um evento específico (i.e. o surgimento de um usuário em cena). Se um usuário aparece em cena, ele tem sua posição identificada, para que em seguida o sensor possa ser calibrado e o usuário seja rastreado. Para cada novo usuário que aparece em cena, ele possui um ID (identificador) para que a aplicação possa fazer referência a cada um individualmente. Para que seus movimentos sejam rastreados, é preciso primeiro detectar sua posição. Este procedimento é feito por meio de uma chamada de *callback*, a qual recebe como parâmetros a capacidade de detectar a pose, o ID daquele usuário, e uma *string* que representa a sua posição. Dessa forma, no caso da detecção de posição ser bem sucedida, é registrada a posição e o processo de detecção de posição é finalizado e é realizada uma requisição à calibração.

Em seguida, é feita a calibração, recebendo como parâmetros a capacidade do esqueleto e o ID do usuário. É na calibração que são medidos e registrados o comprimento dos membros do usuário em cena. Se a calibração completa for executada com sucesso, é feita uma chamada para rastreamento passando como argumento o ID. Caso contrário, o processo de calibração é iniciado novamente e a reinicialização pode ser a partir da detecção de pose do usuário ou da calibração. Se o gerador de calibração exige a detecção de pose, é necessário fazer chamada à função que detecta a pose, para executá-la novamente (passando como argumento o ID do usuário). De outra forma, após realizar a calibração novamente, é passado como argumentos o ID do usuário e o valor "TRUE", este último indica que a calibração vai ser executada novamente para aquele identificador de usuário. Se um usuário sai de cena e as informações referentes ao usuário/cenário que estavam salvas são perdidas, o usuário é declarado como um

usuário perdido e, este tem suas informações excluídas, bem como sua referência. A implementação leva em consideração algum atraso, para que se tenha a certeza de que o usuário realmente saiu do cenário.

Existe uma função (`void SaveCalibration()`) que percorre uma lista de nós existentes, verificando se existem nós calibrados. Em caso afirmativo, eles têm seu ID armazenado em um arquivo. A função (`void LoadCalibration()`) carrega os dados que foram armazenados anteriormente. Esta é muito útil considerando que os desenvolvedores podem armazenar a calibração e, testar posteriormente o aplicativo sem a necessidade de detectar a pose novamente e efetuar a calibração. No caso aqui apresentado, não foi carregado nenhum arquivo (não foi passado como parâmetro).

A função `void glutDisplay()`, é responsável por projetar na tela a representação gráfica dos dados. Por meio do *DepthGenerator* as dimensões *xRes* e *yRes* são acessadas, retornando as dimensões X e Y do *buffer* de profundidade, as quais são necessárias para obter o valor de cada pixel no mapa (de profundidade). Também são incluídos métodos que desenharam o mapa de profundidade, usuários e esqueletos. O gráfico é atualizado apenas quando um nó gerador de usuários tem novos dados.

Por fim, neste arquivo, temos a função principal, onde são definidos o tamanho (em relação à altura/largura) e nome da janela onde serão projetados os gráficos de produção (“User Tracker Viewer”), ocorre a inicialização dos manipuladores de eventos os quais se verifica se a característica de esqueleto (representação gráfica virtual em 2D) de um usuário em cena é suportada fazendo com que o esqueleto se mova de acordo com a movimentação do usuário e suas respectivas dependências (ex. para fazer a calibração, é necessário que exista o suporte da capacidade de calibrar e detectar posição).

No arquivo “*SceneDrawer*” estão dispostos os processadores de evento. Ele é o responsável pela verificação das características (ex. caso seja preciso detectar uma posição do usuário antes de fazer sua calibração, ele registra a detecção de posição, caso contrário, ele não suporta detecção de posição). Esse arquivo é usado também para “desenhar” os membros dos usuários, onde a função `DrawJoint()`, recebe como parâmetros o ID do usuário, as juntas 1 e 2 (por padrão) e traça uma linha reta entre as juntas das articulações especificadas. Verifica também se há o rastreamento dos usuários, por meio de conversões: recebe as coordenadas de posição (real) e as converte em coordenadas projetivas a serem representados em um plano 2D. Além disso, gera um histograma para a criação de uma gradiente de profundidade cena-a-cena, a partir dos parâmetros: longe(escuro)/perto(claro). De acordo com os dados de profundidade, estabelece a textura de cada *pixel* de acordo com o usuário (centro de massa de um usuário). A execução da aplicação é ilustrada na Figura 2.

4.2. Procedimento experimental

Na Figura 2, existem 4 usuários em cena e cada um deles passou pelo processo descrito na seção 4.1 (detecção de pose, calibração e rastreamento). No quadro que é disposto na Figura 3, mostra-se a execução da aplicação. Note que o usuário 1 entrou em cena, teve sua posição detectada e foi calibrado (impresso no terminal). Enquanto o usuário 1 era calibrado, o usuário 2 entrou em cena, teve sua posição detectada e iniciou sua calibração, em paralelo, o usuário 1 tem sua calibração finalizada e seu processo de

rastreamento é inicializado. Quando o usuário 3 entra em cena, ele tem sua posição detectada, é calibrado e rastreado e, logo em seguida, o usuário 2 também tem sua calibração finalizada e então seu processo de rastreamento é inicializado.

Em seguida, entra em cena o usuário 4, este entra e sai da cena duas vezes sua posição é identificada, mas ocorre um erro na tentativa de calibração desse usuário, porém, quando por uma terceira vez, ele entra em cena, tem sua posição detectada e começa o processo de calibração novamente. Enquanto isso, o usuário 5 também entra em cena (e o usuário 3 sai), tem seu posicionamento detectado, mas logo depois, sai de cena. O usuário 4 tem sua calibração executada com sucesso e inicia seu processo de rastreamento. Ocorre falha na tentativa de calibrar o usuário 5 (porque ele saiu de cena).

O usuário 5 entra em cena novamente, tem seu posicionamento detectado, é calibrado com sucesso e inicia-se o processo de rastreamento para o mesmo. Por fim, o usuário 3 (que já saiu de cena há algum tempo) é declarado como perdido. Na janela gráfica “User Tracker Viewer”, na Figura 2, são representados os esqueletos correspondentes aos usuários supracitados, exceto aqueles que saíram de cena, como no caso do usuário 3.

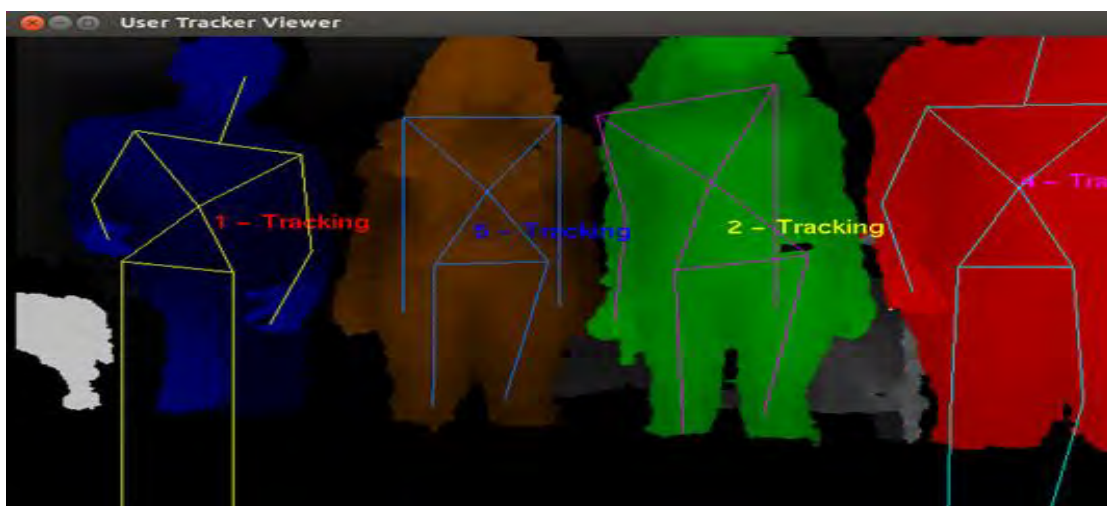


Figura 2. Representação gráfica do resultado da execução da aplicação.

```
lorena@lorena-ubuntu: ~/Documentos/Projeto_OpenNI/OpenNI-Bin-Dev-Linux-x86_64
./0/Samples/Bin/x86-Release$ sudo ./Sample-NiUserTracker
1355314366 New User 1
1355314366 Calibration started for user 1
1355314366 New User 2
1355314366 Calibration started for user 2
1355314366 Calibration complete, start tracking user 1
1355314366 New User 3
1355314366 Calibration started for user 3
1355314368 Calibration complete, start tracking user 3
1355314368 Calibration complete, start tracking user 2
1355314378 New User 4
1355314378 Calibration started for user 4
1355314383 Calibration failed for user 4
1355314383 Calibration started for user 4
1355314388 Calibration failed for user 4
1355314388 Calibration started for user 4
1355314389 New User 5
1355314389 Calibration started for user 5
1355314392 Calibration complete, start tracking user 4
1355314394 Calibration failed for user 5
1355314394 Calibration started for user 5
1355314395 Calibration complete, start tracking user 5
1355314397 Lost user 3
```

Figura 3. Relação do estado da aplicação

5. Conclusão

O presente artigo discutiu acerca do processo de execução de uma aplicação baseada em interação natural. Para tanto, foram apresentados os seus componentes, em termos de bibliotecas, algoritmos e ferramentas físicas, levando em consideração características de rastreamento de movimentos. Nesse contexto, o OpenNI *framework* foi escolhido por oferecer suporte a uma ampla variedade de tipos de dados, o que, em combinação com o sensor Kinect gera informações relacionadas à profundidade, imagem, áudio e movimentos. Com a combinação dessas duas ferramentas, é possível desenvolver uma variedade de aplicações baseadas em interação natural, não só para a área de robótica ou computação gráfica, mas também no contexto de tecnologia social, como por exemplo, no desenvolvimento de aplicações que auxiliem pacientes em fase de reabilitação.

Vale observar que a organização OpenNI, a partir da versão 2.0 não fornecerá suporte às plataformas Mac e Linux, considerando o suporte ao sensor Kinect, uma vez que este é patenteado pela Microsoft. Em contrapartida, a PrimeSense começou a fabricar dispositivos com as configurações similares às do Kinect.

Referências

- Ackerman, E. (2011) “Top 10 Robotic Kinect Hacks”, In: IEEE Spectrum, <http://spectrum.ieee.org/automaton/robotics/diy/top-10-robotic-kinect-hacks>, Março.
- Moore, S. K. (2012) “New Kinect Design Trick: Hand Potter”, In: IEEE Spectrum, <http://spectrum.ieee.org/tech-talk/computing/software/new-kinect-design-trick-handy-potter>, Agosto.
- Rogge, S., Amtsfeld, P., Hentschel, C., Bestle, D. e Maeyer, M. (2012) “Using Gestures To Interactively Modify Turbine Blades In A Virtual Environment”, In: Emergind Signal Processing Applications (ESPA), 2012 IEEE International Conference.
- Wang, Y., Yang, C., Wu, X., Xu, S. e Li, H. (2012) “Kinect Based Dynamic Hand Gesture Recognition Algorithm Research”, In: 4th International Conference on Intelligent Human-Machine Systems and Cybernetics.

SELECTRESS – Sistema Web aplicado ao melhoramento genético vegetal para resistência a fatores ambientais estressantes

Rogério Chaves Barreto¹, Naiara Silva dos Santos¹, Valmir Henrique de Araújo²

¹Centro de Pesquisa e Desenvolvimento de Software– Universidade Estadual do Sudoeste da Bahia (UESB) – Jequié, BA – Brasil

²Universidade Estadual do Sudoeste da Bahia (UESB) – Vit. da Conquista, BA – Brasil
{rogeriocbarreto, naiarassantos, valmirboaideia}@gmail.com

Abstract. Abstract. *This paper aims to describe the SELECTRESS Web System which is being implemented in PHP language programming through the use of resources from other open source libraries. SELECTRESS main goal is to assist students and researches in the area of genetic improvement of plants that are submitted to stressful factors, such as the lack or excess water. Furthermore, this software will provide relevant information on the subject for the entire scientific community.*

Resumo. *Este artigo descreve o Sistema Web SELECTRESS que está sendo implementado na linguagem de programação PHP, utilizando recursos de outras bibliotecas de código aberto. O SELECTRESS tem o objetivo principal de auxiliar estudantes e pesquisadores da área de melhoramento genético de vegetais que são submetidos a fatores estressantes, como, por exemplo, a falta ou o excesso de água. Além disso, tal software disponibilizará informações relevantes sobre o tema para toda a comunidade científica.*

1. Introdução

A área de melhoramento genético tem ganhado importância no setor de pesquisas em produção de animais e vegetais, a fim de obter organismos aperfeiçoados que ao serem disponibilizados no mercado oferecerão aos consumidores produtos de alta qualidade. Estudos da Universidade de São Paulo [USP 2010] acrescentam que, “o melhoramento possibilitou a mecanização das lavouras, o que contribuiu significativamente para um rápido aumento da produtividade agrícola”.

Pesquisadores dessa área têm encontrado dificuldades em seus estudos pela falta de uma ferramenta para automatização de cálculos necessários para acompanhá-los no melhoramento da espécie. Para auxiliar nesse empreendimento, estão sendo desenvolvidos softwares com funções que realizam cálculos, selecionam espécies com genótipos superiores, fazem levantamento estatístico e geram relatórios que auxiliam na tomada de decisão.

O programa GENES – Aplicativo Computacional em Genética e Estatística Experimental – da Universidade Federal de Viçosa (UFV) é um dos softwares que tem por objetivo realizar estatísticas que permitem a análise da diversidade genética de animais e de plantas em geral, gerando informações que são utilizadas por muitos pesquisadores da área de melhoramento genético [Cruz 2008].

Existe também o PROGENIE – Sistema de Auxílio as Decisões de Seleção e Acasalamento em Programas de Melhoramento Genético – que permite o aprimoramento das informações obtidas em pesquisas de melhoramento genético. Esse sistema tem por objetivo auxiliar produtores na seleção de sêmens de bovinos de alta qualidade, para estes serem usados em cruzamentos que resultem em animais com maior ganho genético [Borges, Borges e Ribeiro 2004].

Uma área que carece de estudos é a que está relacionada ao estudo de plantas que são submetidas a fatores ambientais estressantes. Estes fatores podem ser de natureza biótica: as pragas, os fitopatógenos e as ervas daninha; ou abiótica: o calor, a seca, o frio, a salinidade, o desequilíbrio nutricional, o alagamento e os metais pesados. Tais fatores são condições que influenciam no crescimento, na produtividade e na reprodução das plantas, ocasionando prejuízos tanto para os consumidores como também para os produtores. Levanta-se, então o questionamento: como desenvolver um software que auxilie os pesquisadores que muitas vezes têm que realizar cálculos manualmente que demandam muito tempo e possuem baixa confiabilidade nos resultados obtidos? Para solucionar esta problemática, o software SELECTRESS está sendo criado com o objetivo de ajudar os melhoristas vegetais. Para tal fim, ele executará cálculos de índices que auxiliam os pesquisadores a determinarem qual a suscetibilidade dos vegetais a determinadas condições estressantes. O que diferencia o SELECTRESS dos softwares existentes é a especificidade da abordagem referente ao melhoramento genético de plantas que vivem nas condições estressantes citadas acima. A realização de tais cálculos por esse sistema agilizará e automatizará a obtenção dos resultados, o que tornará os estudos menos dispendiosos e mais confiáveis. Além da realização dos cálculos e da automatização do sistema, o SELECTRESS exibirá informações sobre o estresse ambiental ao qual a planta está submetida.

2. Modelagem e Implementação

O SELECTRESS está sendo desenvolvido na plataforma Web em busca de torná-lo mais acessível em diferentes plataformas de sistemas operacionais e poder acompanhar assim a tendência da virtualização dos aplicativos computacionais. A virtualização é abordada por Pierre Lévy (2007) ao afirmar que “um movimento geral de virtualização afeta hoje não apenas a informação e a comunicação, mas também os corpos, o funcionamento econômico, os quadros coletivos da sensibilidade ou o exercício da inteligência”.

A implementação do sistema proposto tem como base a utilização de uma linguagem de programação Web, para tornar o conteúdo disponível a toda sociedade,

principalmente aos pesquisadores da área de melhoramento genético vegetal, servindo como mais uma ferramenta de apoio às suas pesquisas.

O método de desenvolvimento utilizado é o Scrum Solo, uma adaptação do método original Scrum. Ele é dividido em artefatos como o Product backlog, que é uma lista de itens a serem desenvolvidos para um software de acordo com as prioridades do cliente, o Sprint backlog que são tarefas concretas necessárias para a conclusão de partes do Product Backlog. O Sprint é a definição de quanto trabalho será necessário para implementar um item de backlog, que pode ser uma semana ou até um mês [Schwaber 2009]. Neste método adaptado foi utilizado sprints de uma semana, em que não haverá reuniões diárias, apenas reuniões semanais com a equipe de validação, que são os possíveis usuários do sistema.

A criação do SELECTRESS iniciou-se com o levantamento de requisitos que foi realizado através de uma entrevista aberta junto a um pesquisador da área de melhoramento genético de vegetais, com o intuito de obter dados relevantes para a criação do Product backlog. Após esta etapa, foram criados os Sprints backlog a fim de definir os passos necessários para a conclusão dos itens do Product Backlog. Os Sprints constituíram a modelagem de cada Sprint backlog utilizando casos de usos e diagramas, como por exemplo, o de classes e o de atividades, além da implementação do backlog.

O Sprint mais relevante foi o do desenvolvimento da calculadora para realizar os cálculos dos índices que influenciam no desenvolvimento da planta, o cerne do sistema. Para tal fim foi utilizado funções construídas com a linguagem JavaScript. Um exemplo simples de cálculo que o sistema realizará é mostrado a seguir:

PM – PRODUTIVIDADE MÉDIA PSFE = Produção do Genótipo Específico em Condições Não–Estressantes PCFE = Produção do Genótipo Específico em Condições Estressantes	
$PM = \frac{PSFE + PCFE}{2}$	[Rosielle e Hamblin 1978]

Figura 1. Fórmula de Produtividade Média

3. Resultados Esperados

A partir da modelagem e da criação do software SELECTRESS, espera-se que as informações relativas aos fatores que influenciam no crescimento e na produtividade das plantas cultivadas em condições estressantes sejam fornecidas e disponibilizadas para toda a comunidade científica e estudantes de áreas afins, otimizando e auxiliando cientistas na escolha de progênies mais adaptadas às condições as quais estão sujeitas.

Os cálculos dos índices que interferem na evolução dos vegetais foram automatizados, para agilizar os trabalhos dos pesquisadores e proporcionar uma maior confiabilidade dos resultados obtidos, servindo de apoio na escolha das plantas mais adaptadas às condições estressantes, como por exemplo, o calor, a deficiência mineral e outros, que o homem e a natureza as submetem interferindo na vida útil da planta.

O SELECTRESS será submetido a testes com pesquisadores e estudantes do setor de melhoramento genético vegetal. Além disso, os usuários terão um espaço reservado para deixarem seus comentários, proporcionando, portanto, um feedback sobre a funcionalidade do sistema. Isso fará com que relatórios e estatísticas de uso possam ser gerados. Desta forma, será possível identificar se o sistema está realmente automatizando os cálculos e auxiliando melhoristas vegetais.

4. Conclusão

Devido ao aumento da população mundial e da consequente demanda por alimentos de alta qualidade, pesquisadores e cientistas têm investido na área de melhoramento genético com o intuito de selecionar indivíduos geneticamente superiores. Para auxiliar nesse empreendimento é de fundamental importância o desenvolvimento de softwares que agilizam a realização de etapas do estudo fornecendo uma maior confiabilidade aos resultados obtidos. Um desses softwares é o SELECTRESS.

Por estar sendo desenvolvido na plataforma Web, o SELECTRESS poderá ser acessado em qualquer sistema operacional e em qualquer localização geográfica sem ter alterações na sua segurança, fazendo com que o tema abordado seja amplamente difundido.

Referências

- Borges, M., Borges, G. S. N. e Ribeiro, H. G. (2004) “PROGENIE – Sistema de auxílio às decisões de seleção e acasalamento em programas de melhoramento genético.”, V Simpósio da Sociedade Brasileira de Melhoramento Animal –, Pirassununga – SP.
- Cruz, C. D. (2008) “Programa Genes: diversidade genética.” Editora UFV, 1ª Edição, Viçosa – Minas Gerais.
- Lévy, Pierre (2007) “O que é o virtual? / Pierre Lévy; tradução de Paulo Neves”, Editora 34, São Paulo.
- Pressman, R. S. (2006) “*Engenharia de Software*”, McGraw-Hill, 6ª edição.
- Rosielle, A.A., and Hamblin, J. 1981. Theoretical Aspects of Selection for Yield in Stress and Non-stress Environments. Crop Sci., 21: 943-946.
- Sommerville, I. (2007) “*Engenharia de Software*”, Editora Pearson, 8ª edição.
- Schwaber, K. (2009). Agile project management with Scrum. Microsoft Press.
- Universidade de São Paulo, (2010) “Melhoramento Genético”. Instituto de Biociências. Web: http://felix.ib.usp.br/bib141/melhoramento_genetico.pdf

Modelagem Multifractal β MWM Adaptativa para Tráfego de Redes de Computadores

Bruno Henrique Pereira Gonçalves¹, Flávio Henrique Teles Vieira¹,
Victor Hugo Teles Costa¹

¹Escola de Engenharia Elétrica, Mecânica e de Computação
Universidade Federal de Goiás (UFG) – Goiânia, GO – Brasil

brunohpg@gmail.com, flavio@eee.ufg.br, victor@vcosta.com.br

Abstract. *In this paper, we propose an adaptive model based on the Multifractal Wavelet Model (MWM). We show how the parameters of the β MWM, that is a particular form of the MWM, can be estimated in real-time through the proposed approach that uses the Haar wavelet. We verify the performance of the modeling comparing the statistics of real traffic series with statistics of synthetic series generated by the model. We show that the proposed adaptive modeling provides similar results to those of the original modeling.*

Resumo. *Neste artigo propomos uma modelagem adaptativa baseada no Modelo Wavelet Multifractal (MWM-Multifractal Wavelet Model). Mostramos como os parâmetros do β MWM, que é um caso particular do MWM, podem ser estimados em tempo real através da abordagem proposta onde utilizamos a Wavelet de Haar. Verificamos o desempenho da modelagem comparando as estatísticas de séries reais de tráfego de redes de computadores com estatísticas de séries sintéticas geradas pelo modelo. Mostramos que a modelagem adaptativa proposta apresenta resultados similares aos da modelagem original.*

1. Introdução

O desempenho de redes de comunicações é alvo de pesquisas constantes, e muitos são os trabalhos propondo modelos para descrição do tráfego de redes [1][2][3][4]. Técnicas de análise que proporcionam melhor compreensão do comportamento do tráfego de redes são muito importantes na concepção e otimização de redes de comunicações [5]. Uma vez obtido um modelo adequado para o tráfego, profissionais da área de redes podem prever o desempenho da rede por meio de técnicas analíticas ou por meio de simulação e assim proporem mecanismos mais eficientes para provisão de QoS (Qualidade de Serviço) e controle dos fluxos que trafegam pela rede.

Algumas características do tráfego de redes de computadores não são consideradas por modelos de tráfego mais simplificados, como modelos Markovianos. Muitos estudos revelam a alta variabilidade de fluxos de redes, ou seja, o tráfego contém rajadas em várias escalas de tempo, em contraste a suposição de que rajadas só existem em escalas curtas de tempo [6][7]. Foi mostrado que estas incidências de rajadas multiescalas tem um impacto significativo no desempenho das redes [6][7][8]. Modelos mais realistas foram propostos, inicialmente caracterizando a autossimilaridade do tráfego[9]. O termo autossimilaridade se refere normalmente a processos assintoticamente autossimilares de

segunda ordem ou monofractais [10]. O parâmetro de Hurst mede o grau de autossimilaridade do processo. No entanto, para muitos processos de tráfego de rede, a energia em escala dos coeficientes *wavelet* ou os gráficos variância-tempo normalmente não demonstram comportamento linear. Muitos destes processos têm comportamento fractal com parâmetro de Hurst variado em diferentes escalas de tempo pequenas [11], ou seja, são multifractais. Introduzido por Mandelbrot [12], os modelos multifractais foram aplicados amplamente a várias áreas científicas. No campo de tráfego de redes, vários modelos foram propostos [1][3][4][13][14]. O MWM (*Multifractal Wavelet Model*)[14] é um modelo multifractal que se destaca por aplicar a transformada *wavelet* à série de tráfego e apresentar várias alternativas de modelagem.

Muitas técnicas de modelagem de tráfego encontradas na literatura partem de um conhecimento prévio das características do tráfego. A modelagem do tráfego em tempo real que se adapta às variações de características do fluxo é muito interessante para tomada de decisões adequadas relativas à concepção e gestão de redes.

Neste trabalho, propomos uma modelagem multifractal adaptativa baseada no modelo β MWM - derivada da modelagem MWM - e apresentamos o desempenho da modelagem em função do número de amostras fornecidas ao modelo. A estimação adaptativa dos parâmetros do modelo permite que estes parâmetros sejam adaptados de forma a acompanhar as mudanças do tráfego a ser modelado. Com isso, diferente de um modelo estático, temos uma atualização contínua do modelo, que pode ser utilizado em aplicações que envolvam mudanças no comportamento do tráfego.

O artigo está organizado da seguinte forma: na seção 2, abordamos os princípios do funcionamento do modelo β MWM; na seção 3 apresentamos nossa proposta de modelo β MWM adaptativo; na seção 4 apresentamos o desempenho da modelagem para algumas séries de tráfego reais; na seção 5 fazemos algumas considerações finais sobre o trabalho.

2. β MWM

O *Multifractal Wavelet Model* (MWM) é um modelo multifractal com grande destaque na modelagem de tráfego de redes [14][15]. Ele é baseado em uma cascata multiplicativa no domínio *wavelet*. A transformada *wavelet*[16][17] discreta é usada neste modelo devido a sua capacidade de representação multiescala de sinais. O MWM apresenta mais de uma modelagem para os coeficientes *wavelet* e coeficientes de escala, gerados pela transformada *wavelet*. Uma dessas modelagem é o β MWM.

O processo de modelagem do β MWM realiza a transformada discreta de *wavelet* de Haar para um número fixo de camadas, J , da cascata multiplicativa binomial[4] para a série completa em uma única etapa. A partir dos coeficientes *wavelet* ($W_{j,i}$) e coeficientes de escala ($U_{j,i}$) gerados, por camada $j \quad \forall \quad 0 \leq j \leq J - 1$, os parâmetros MWM são estimados.

Para se compreender o modelo MWM, devemos citar alguns conceitos da transformada *wavelet*. A transformada *wavelet* discreta é usada para representação multiescala de sinais da seguinte forma:

$$f(t) = \sum_k U_{J_0,k} \phi_{J_0,k}(t) + \sum_{j=J_0}^{\infty} \sum_k W_{j,k} \varphi_{j,k}(t) \quad (1)$$

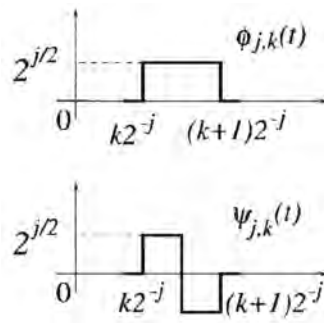


Figura 1. Funções escala $\phi_{j,k}(t)$ e *wavelet* de Haar $\varphi_{j,k}(t)$

onde $W_{j,k}$ e $U_{j,k}$ são respectivamente os coeficientes *wavelet* e de escala, dados por:

$$W_{j,k} = \int f(t) \varphi_{j,k}(t) dt \quad (2)$$

e

$$U_{j,k} = \int f(t) \phi_{j,k}(t) dt \quad (3)$$

A Figura 1 exibe as funções escala $\phi_{j,k}(t)$ e *wavelet* de Haar $\varphi_{j,k}(t)$ utilizadas na representação multiescala do sinal. Pode-se demonstrar que os coeficientes de escala dados pela equação (3) podem ser recursivamente calculados utilizando a *wavelet* de Haar $\varphi_{j,k}(t)$ através das seguintes equações:

$$U_{j,2k} = 2^{-1/2} (U_{j-1,k} + W_{j-1,k}) \quad (4)$$

$$U_{j,2k+1} = 2^{-1/2} (U_{j-1,k} - W_{j-1,k}) \quad (5)$$

Este processo recursivo é repetido até que se atinja a resolução desejada ou equivalentemente, até que se obtenha o número desejado de amostras, formando uma árvore binária de coeficientes de escala. A Figura 2 apresenta a árvore binária formada para obtenção dos coeficientes de escala representada pelas equações (4) e (5). No modelo MWM, a fim de assegurar a não-negatividade da série de tráfego sintético, determinadas condições devem ser impostas a seus coeficientes *wavelet* e de escala. Os coeficientes $U_{j,k}$ representam a média local do processo em escalas e deslocamentos de tempo diferentes. A condição $X(t) \geq 0, \forall t$, impõe que, $U_{j,2k+1} \geq 0, \forall j, k$. Impondo a condição $U_{j,k} \geq 0, \forall j, k$, pode-se afirmar que $|W_{j,k}| \leq U_{j,k}, \forall j, k$. Os coeficientes *wavelet* são gerados a partir da equação:

$$W_{j,k} = U_{j,k} A_{j,k} \quad (6)$$

onde $A_{j,k}$ é uma variável aleatória cujo valor está em $[-1, 1]$. Além disso, supõe-se algumas condições para esta variável: os multiplicadores $A_{j,k}$ são independentes e identicamente distribuídos (i.i.d) dentro de cada escala, são também independentes de $U_{j,k}$ e simétricos em torno de zero. No caso particular do β MWM, os multiplicadores $A_{j,k}$ são modelados segundo uma distribuição beta simétrica[18]. A função densidade de probabilidade da distribuição beta simétrica é dada por:

$$f(x) = \frac{(1+x)^{p-1}(1-x)^{p-1}}{B(p,p)2^{2p-1}} \quad (7)$$

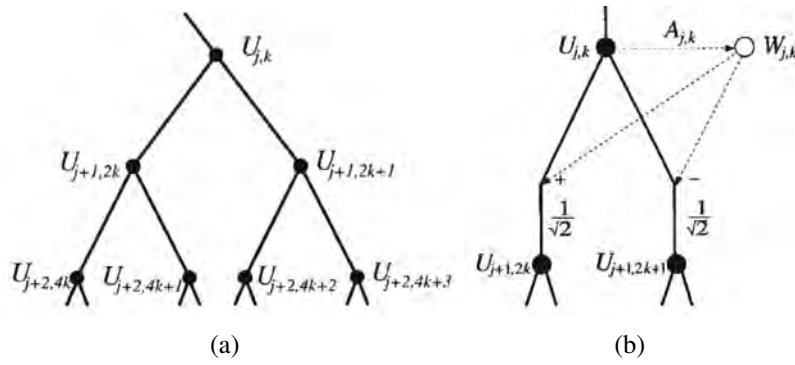


Figura 2. Árvore binária dos coeficientes de escala

onde $B(\cdot, \cdot)$ é a função beta e p é o parâmetro que determina a forma da distribuição. Os multiplicadores $A_{j,k}$ são escolhidos de forma a controlar as energias dos coeficientes *wavelet*. Assim, obtém-se as seguintes relações [14]:

$$n_j = \frac{E(W_{j-1,k}^2)}{E(W_{j,k}^2)} = \frac{2p_j + 1}{p_{j-1} + 1} \quad (8)$$

e

$$(2p_0 + 1)E(W_{0,0}^2) = E(U_{0,0}^2) \quad (9)$$

Pode-se notar que p_j é usado para capturar o decaimento da energia dos coeficientes *wavelet* em escala, n_j . O coeficiente de escala $U_{0,0}$ na maior escala (mais fina) é modelado como sendo uma variável aleatória normal com média, μ_c , e a variância, σ_c^2 , iguais aos dos coeficientes de escala dos fluxos de tráfego reais.

O β MWM relaciona o decaimento de energia dos coeficientes *wavelet* n_j por camada j com os valores dos parâmetros p_j das distribuições beta simétrica, utilizadas para modelar os multiplicadores $A_{j,k}$. Os parâmetros p_k são estimados, recursivamente, por:

$$p_j = \frac{n_j}{2} (p_{j-1} + 1) - \frac{1}{2} \quad (10)$$

Considerando um tempo discreto k , tem-se que o processo discreto de tráfego MWM $X[k]$ é obtido pelos coeficientes de escala $U_{j,k}$ na escala mais fina j , da seguinte forma:

$$X[k] = 2^{-\frac{j}{2}} U_{j,k} \quad (11)$$

O MWM pode precisamente modelar a dependência de longa duração presente nos dados de tráfego assim como capturar outras características multifractais. Entretanto, uma desvantagem do MWM é o número de parâmetros a serem estimados, que para isso, faz uso de toda série de tráfego. Dessa forma, o MWM se torna inadequado a aplicações em tempo real.

3. β MWM Adaptativo

Nesta seção, propomos um modelo β MWM Adaptativo. Ao invés do processamento de todos os dados da série de tráfego em uma única etapa, propomos o processamento iterativo em janelas de tamanho fixo de 2^J amostras, onde J é o número de camadas da cascata

modeladas. Apenas algumas variáveis são armazenados no processo de modelagem, não havendo a necessidade de guardar uma grande quantidade de dados sobre o fluxo. Note que as séries de tráfego neste trabalho representam quantidade de bytes em determinados instantes de tempo.

A modelagem proposta é feita da seguinte forma: Primeiramente as variáveis do modelo são iniciadas, ou seja, fazemos o segundo momento dos coeficientes *wavelet*, $E[W_{j,k}^2](0) = 0$; média e variância dos coeficientes de escala $\mu_c(0) = 0$ e $\sigma_c^2(0) = 0$; contador de janela $n = 0$. E depois, a cada janela não sobreposta de 2^J amostras:

Algoritmo 1: Algoritmo para Estimativa Adaptativa dos Parâmetros do β MWM

1º passo Realiza-se a transformada de Haar na janela de dados. A transformada de Haar em cada janela de 2^J amostras gera 2^j coeficientes *wavelet* - nomeados de $\tilde{W}_{j,k}$ - por camada j e um coeficiente de escala - nomeado de $\tilde{U}_{0,0}$ - na camada $j = 0$;

2º passo Atualiza-se o segundo momento $E[W_{j,k}^2]$ dos coeficientes *wavelet* através da equação:

$$E[W_{j,k}^2](n+1) = E[W_{j,k}^2](n) \left(\frac{n}{n+1} \right) + \frac{\sum_{i=0}^{2^j-1} \tilde{W}_{j,i}^2}{(n+1)2^j} \quad (12)$$

3º passo As taxas de energia n_j são recalculadas segundo a equação (8) e os parâmetros p_j são recalculados segundo a equação (10).

4º passo As estatísticas dos coeficientes de escala são atualizados segundo as equações:

$$\mu_c(n+1) = \mu_c(n) \left(\frac{n}{n+1} \right) + \frac{\tilde{U}_{0,0}}{n+1} \quad (13)$$

$$\sigma_c^2(n+1) = \left(\sigma_c^2(n) + (\mu_c(n))^2 \right) \left(\frac{n}{n+1} \right) - (\mu_c(n+1))^2 + \frac{(\tilde{U}_{0,0})^2}{n+1} \quad (14)$$

O processo é repetido a cada nova janela de dados de 2^J amostras, incrementando o valor da variável n em 1. Assim, com o algoritmo adaptativo proposto obtém-se os parâmetros do modelo β MWM que são: p_j s, μ_c e σ_c^2 .

4. Resultados

Para avaliar o desempenho da modelagem proposta, modelamos as séries de tráfego Internet/TCP-IP reais dec-pkt-3 e dec-pkt-4¹, ambas agregadas em intervalos de 512ms. Para a série dec-pkt-3, além dos coeficientes de escala, três camadas da cascata multiplicativa foram modeladas, ou seja, $J = 3$; e para a série dec-pkt-4, quatro camadas, $J = 4$. As simulações foram realizadas utilizando-se um computador com processador Intel Core i5 e 4GB de memória RAM.

A Figura 3(a) apresenta os valores dos parâmetros beta simétrico, p_j s, para cada camada j , $j = 0, 1, 2$, estimados para a série dec-pkt-3-512ms em função do número de amostras da série. A série dec-pkt-3-512ms corresponde aos valores de quantidade de

¹<http://ita.ee.lbl.gov/html/contrib/DEC-PKT.html>

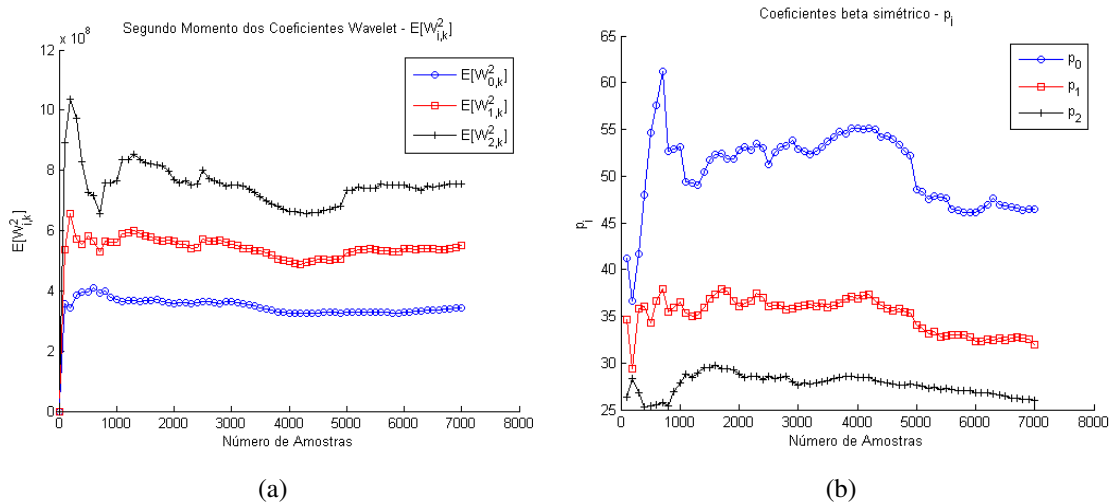


Figura 3. Segundo momento dos coeficientes *wavelet* e parâmetros beta simétrico estimados para a série dec-pkt-3-512ms

bytes na escala de tempo de 512ms da série TCP-IP dec-pkt-3. A Figura 3(b) apresenta o segundo momento dos coeficientes *wavelet*, $E[W_{j,k}^2]$, para cada camada j , em função do número de amostras. As Figuras 4(a) e 4(b) mostram, respectivamente, a evolução dos valores dos parâmetros média, μ_c , e variância, σ_c^2 , dos coeficientes de escala gerados pela transformada *wavelet* para a série dec-pkt-3-512ms pelo número de amostras da série. Pode-se observar pela Figura 4 que os valores de média, μ_c , e variância, σ_c^2 , dos coeficientes de escala tendem à medida que o número de amostras aumenta aos valores obtidos com o MWM tradicional (*on batch*) cujos valores são $\mu_c = 2.6161 \cdot 10^5$ e $\sigma_c^2 = 2.3783 \cdot 10^9$.

Comparamos os momentos das séries reais com os momentos de séries sintéticas geradas a partir da modelagem em função do número de amostras. As curvas de momentos geradas pela modelagem β MWM não adaptativa se sobrepõe as curvas de momentos geradas pela modelagem proposta para o mesmo número de amostras. Essas comparações são apresentadas na Figura 5, para a série dec-pkt-4-512ms. O processo de síntese de uma série sintética a partir do modelo MWM está descrito em [14].

As figuras mostram que a modelagem β MWM adaptativa consegue modelar estatísticas de várias ordens com grande precisão. A maior parte dos parâmetros tendem a se estabilizarem com o aumento do número de amostras, mostrando que o algoritmo conseguiu extrair as principais características da série antes de terminar o processamento de todas as amostras.

5. Conclusão

Associando estimação adaptativa de parâmetros ao MWM conseguimos modelar estatísticas de várias ordens do tráfego de rede em tempo real, adaptando os parâmetros do modelo às variações do fluxo de tráfego. Os resultados apresentados mostram que o modelo adaptativo descreve tão bem quanto ao MWM *on batch* as estatísticas dos fluxos de tráfego reais considerados.

Devido a tal resultado de desempenho, o modelo adaptativo proposto pode ser

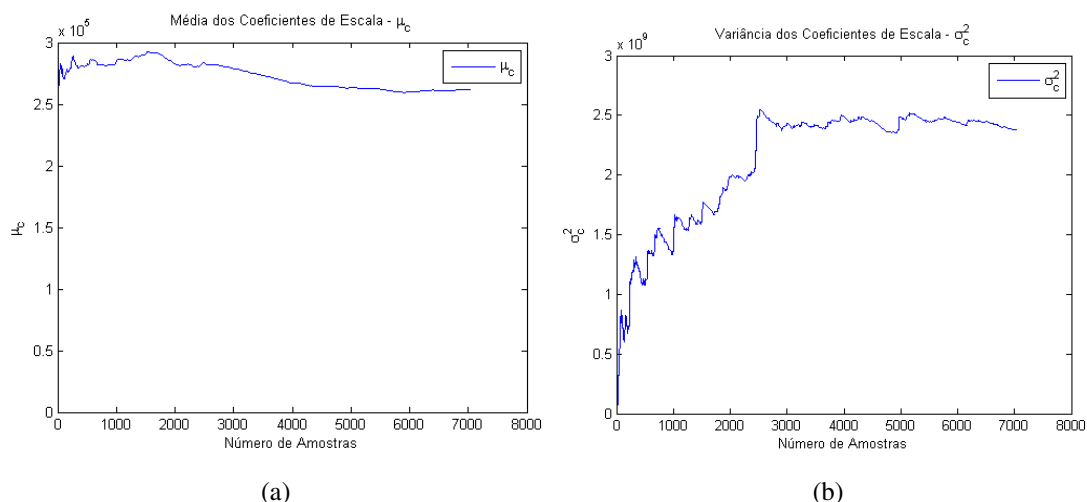


Figura 4. Parâmetros Média e Variância dos Coeficientes de Escala da série dec-pkt-3-512ms

aplicado a várias técnicas de controle do fluxo de tráfego de redes de computadores, aprimorando tais técnicas.

Referências

- [1] Krishna, M. P., Gadre, V. M. e Dessay, U. B. (2003) "Multifractal based network traffic modeling". Kluwer Academic Publishers.
- [2] Lee, I. W. C. e Fapojuwo, A. O. (2005) "Stochastic processes for computer network traffic modelling". Computer Communications, vol. 29, pp.1-23.
- [3] Vieira, F. H. T. e Ling, L. L. (2008) "Modelagem de Tráfego de Redes Utilizando Cascata Multifractal Generalizada". RITA, Vol.15, No. 2.
- [4] Rocha, F. G. C. e Vieira, F. H. T. (2009) "Modelagem de tráfego de vídeo MPEG-4 utilizando cascata multifractal com distribuição autorregressiva dos multiplicadores". I2TS, Florianópolis, SC.
- [5] Frost, V. and Melamed, B. (1994) "Traffic modeling for telecommunications networks". IEEE Communications Magazine 32:3, 70-81.
- [6] Leland, W. E., Taqqu, M. S., Willinger, W., Wilson, D. V. (1994) "On the self-similar nature of Ethernet traffic" (extended version). IEEE/ACM Transactions on Networking, v. 2, n. 1, pp.1-15.
- [7] Paxson, V., Floyd, S. (1995) "Wide-area traffic: The failure of Poisson modeling". IEEE/ACM Transactions on Networking, v.3, n.3, pp.226-244.
- [8] Erramilli, A., Narayan, O., Willinger, W. (1996) "Experimental Queueing Analysis with Long-Range Dependent Packet Traffic". IEEE/ACM Trans. on Net., v. 4, n.2, Abril.
- [9] Liebeherr, J., Patek, S. D., Burchard, A. (2003) "Statistical per-flow service bounds in a network with aggregate provisioning". IEEE INFOCOM., v. 3, 30 Março a 3 Abril, pp.1680-1690.
- [10] Park, K., Willinger, W. (2000) "Self-similar Network Traffic and Performance Evaluation", John Wiley and Sons, New York.

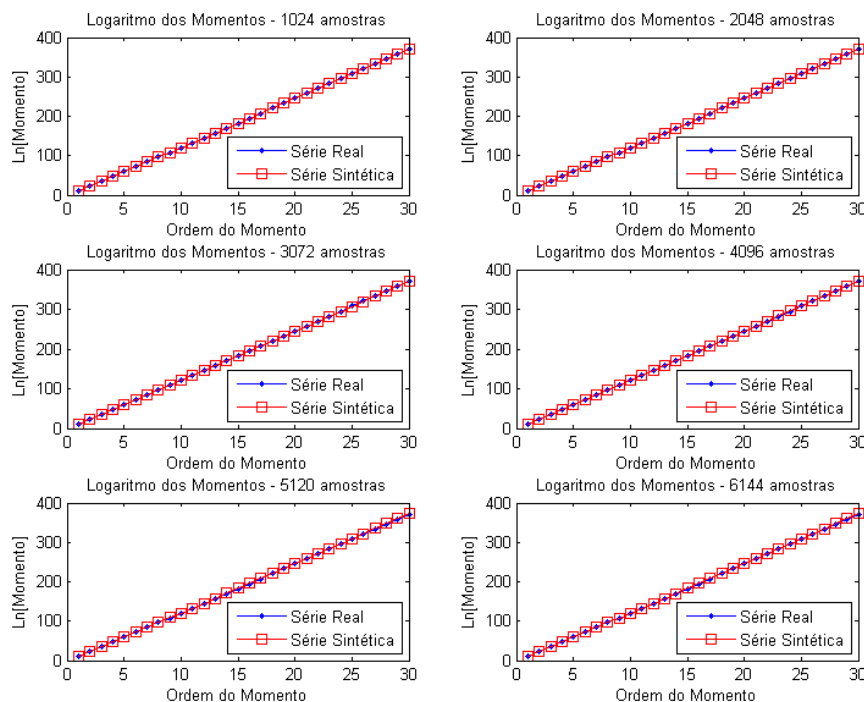


Figura 5. Comparação entre momentos da série dec-pkt-4-512ms e momentos da série a partir do modelo

- [11] Li, Q., Mills, D.(1999) “Investigating the Scaling Behavior, Crossover and Antipersistence of Internet Packet Delay Dynamics”. Globecom 3,pp. 1843-1852.
- [12] A. Fisher, L. Calvet, and B. B. Mandelbrot (1997) “Multifractality of Deutschmark/US Dollar Exchanges Rates”. Yale University.
- [13] S. Molnár, T. D. Dang, I. Maricza (2002) “On the queue tail asymptotics for general multifractal traffic”, In Proc., IFIP Networking 2002, Pisa, Italia.
- [14] Riedi, R. H., Crouse, M. S., Ribeiro, V. J. e Baraniuk, R. G. (1999) “A multifractal wavelet model with application to network traffic”. IEEE Trans. on Information Theory, vol. 45, no.3, pp. 992-1018.
- [15] Ribeiro, V. J., Riedi, R. H., Crouse, M. S., Baraniuk, R. G. (2000) “Multiscale queueing analysis of long-range dependent traffic”. Proc. IEEE Infocom.
- [16] Vieira, F. H. T., Costa, V. H. T., Gonçalves, B. H. P. (2013) “Neural Network Based Approaches for Network Traffic Prediction”. Artificial Intelligence, Evolutionary Computing and Metaheuristics, pp. 657-684
- [17] Chui, C. K. (1992) “An Introduction to Wavelets”. San Diego: Academic.
- [18] Spiegel, M. R.; Liu, J. (2011) “Manual de Fórmulas e Tabelas Matemáticas”. Col. Schaum - 3ª Ed, Makron Books.

Um Sistema Móvel de Compras Rápido e Seguro via QR Code e Vitrine Virtual

Eduardo R. Costa, Karin S. Komati

Coordenadoria de Informática, Instituto Federal do Espírito Santo (IFES) Campus Serra
Rodovia ES-010, Km 6.5 – 29.164-231 – Serra – ES – Brasil

erigamonte@gmail.com, kkomati@ifes.edu.br

Abstract. *This work aims to present a sales system for mobile devices that is safe and agile. Particularly, the issue of safety was a major concern of this work. Our proposed solution is based on the use of QR Code to direct the user to the web store through their mobile devices; a strategy that we call virtual showcase. Unfortunately, this strategy potentially exposes the user of our system to different kinds of attack. We propose the use of symmetric key cryptography as a way to prevent these attacks. In this way we expect to guarantee, by the use of cryptography, that the information read from the QR Code is indeed valid.*

Resumo. *O objetivo deste trabalho é apresentar um sistema de compras para dispositivos móveis que seja rápido e seguro. Em particular, o quesito segurança foi a maior preocupação do trabalho. A solução proposta se baseia na utilização de QR Code para direcionar os usuários, através de seus dispositivos móveis, para o site de venda, estratégia que chamamos de vitrine virtual. Entretanto, esta estratégia expõe os usuários do sistema a diversos tipos de ataques. Propomos o uso de criptografia baseada em chaves simétricas como uma forma de evitar tais ataques. Desta forma buscamos garantir, através do uso de criptografia, que a informação lida do QR Code seja válida.*

1. Introdução

O comércio eletrônico *mobile* (*m-commerce*) está ganhando cada vez mais espaço. Segundo o instituto de pesquisas IDC [G1, 2012], até o ano de 2017 as compras feitas pelo celular chegarão a um montante de US\$ 1 trilhão. A maior parte das compras inclui músicas, filmes e aplicativos, e recentemente, mercadorias, como: eletrônicos, roupas e alimentos. No entanto, muitas vezes a navegação em diversas páginas pode ser comprometida pelo tamanho da tela, ou as pessoas não possuem a destreza de digitar textos extensos nestes aparelhos. Para facilitar e agilizar o uso dos celulares, temos a tecnologia denominada de QR Code (*Quick Response Code*), que foi desenvolvido pela Denso Wave, lançado em 1994 e padronizado com o ISO/IEC 18004 [Denso Wave Corporation, 2012].

O QR Code é uma imagem, um código de barras 2D, em duas cores, que pode ser lido e interpretado pelos *smartphones*. Quando a informação contida pelo QR Code for uma URL, direciona-se automaticamente para o endereço fornecido na internet, tornando o acesso fácil e rápido [Komati, Costa e Andrade, 2012]. Uma característica visível é a presença de três quadrados nos vértices do código, chamados de padrões de localização. Esses padrões de detecção de posição garantem a leitura estável, evitando os efeitos negativos da interferência de fundo e possibilitando a leitura de 360° em alta

velocidade. Enquanto os códigos de barras convencionais armazenam no máximo 107 dígitos, o QR Code pode armazenar até 7089 dígitos.

Nesse contexto, o objetivo deste trabalho é apresentar um sistema de compras, o **Mobile Market**, rápido e seguro, utilizando a tecnologia QR Code e a ideia de vitrine virtual. Uma vitrine virtual pode ser uma revista, folheto ou encarte, que em geral é distribuído gratuitamente ou; pode estar na forma de cartazes dispostos em lugares públicos. O importante é que cada produto tenha um QR Code associado. A Figura 1 mostra o funcionamento do processo de compra: a vitrine virtual nesse caso é um encarte, onde o usuário aponta a câmera do celular para o símbolo QR Code do produto que deseja comprar. O aplicativo faz a leitura do código, decodifica-o, e comunica-se com o site, e vai enchendo o carrinho de compras.



Figura 1. Funcionamento de leitura do módulo Mobile, mostrando a leitura do QR Code e o acesso do produto no sistema da loja.

Já existem soluções similares no mercado como a da Multinacional Tesco [Estadão, 2012]. Esta empresa modificou a estrutura do QR Code original, pois os códigos apresentados mostram um padrão com 4 estruturas nos cantos do QR Code ao invés das 3 comumente usadas, programas genéricos de leitura de código QR Code não conseguem ler o conteúdo destes códigos. A grande motivação do uso de um código modificado se deve aos problemas quanto à segurança. Seres humanos não conseguem diferenciar entre imagens de QR Codes válidos ou aquelas que foram maliciosamente alterados. Um incidente aconteceu na Rússia [Wasserman, 2011], onde um QR Code adulterado enganou os consumidores que pensavam estar baixando um aplicativo Android chamado Jimm. O código continha um *malware* que enviou códigos SMS para um número de telefone que cobrava por cada mensagem enviada.

Assim, um diferencial deste trabalho é a inclusão da informação gravada no QR Code de forma cifrada. Diferente da solução da Tesco que alterou a estrutura do QR Code, este trabalho propõe criptografar a informação e manter a estrutura padrão do QR Code. Com o uso de criptografia de chave simétrica, garante-se que o QR Code lido é exatamente o mesmo QR Code gerado pela mesma empresa que está vendendo os produtos.

2. Solução Proposta: Mobile Market

O sistema terá dois módulos: a parte **Desktop**, e o aplicativo **Mobile**. A parte **Desktop** é

responsável pela criação do QR Code para estar associado a cada produto na vitrine virtual. O QR Code deverá conter as seguintes informações: URL do produto no *site* para localizar o produto desejado no *site* da loja, a data de validade do preço que se encontra na vitrine virtual e o preço tal qual na vitrine.

Uma das funcionalidades do módulo **Desktop** é a encriptação das informações e criação dos QR Codes, que serão associados na vitrine para a identificação no sistema do supermercado. Esse módulo foi feito em Java usando a API QRGen [Gullaksen, 2012] para a geração do QR Code. Os dados do QR Code estarão cifrados usando a técnica de criptografia de chave simétrica [cert.br, 2012], conforme a parte esquerda da Figura 2, denominada de “Geração”.

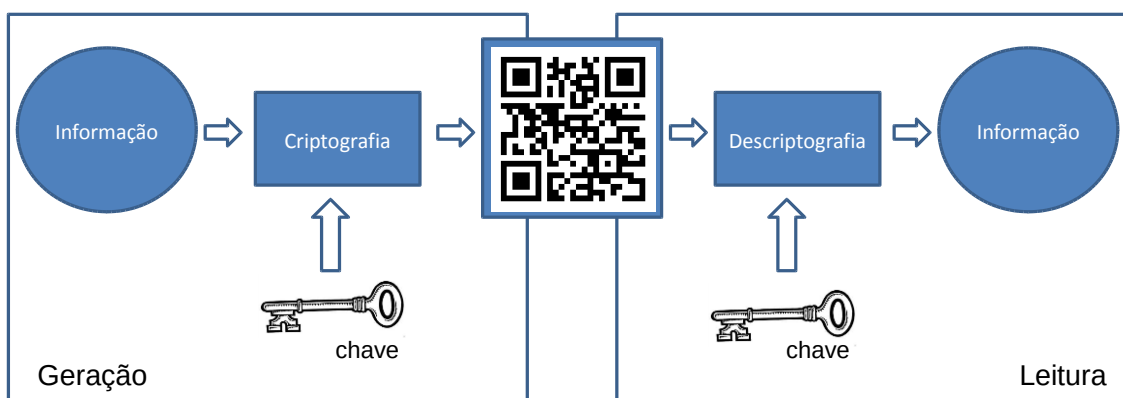


Figura 2. Fluxograma que mostra o passo a passo do processo de encriptação e descriptação da informação contida no QR Code.

O aplicativo **Mobile** faz a leitura do código, decodifica-o, e comunica-se com o site, conseguindo as informações disponibilidade (ou quantidade em estoque), uma pequena imagem do produto e o preço atual do *site*. De forma fácil, o usuário poderá identificar quantos produtos ele deseja comprar ou mesmo se deseja retirar algum item do carrinho de compras. Ao finalizar o pedido, o usuário informará o endereço de entrega e uma forma de pagamento. Se a data de validade do preço, constante no QR Code, estiver expirada, o sistema irá alertar o usuário de que o produto teve seu preço reajustado e qual é o novo preço de acordo com o *site*.

Como é o aplicativo **Mobile** que fará a decodificação da informação gravada no QR Code e os dados no QR Code estarão cifrados, o usuário só conseguirá fazer as compras se usar o módulo **Mobile** fornecido, conforme a parte direita da Figura 2, denominada de “Leitura”. Há o inconveniente de se fazer o *download* deste aplicativo no celular, mas será apenas uma vez.

O módulo **Mobile** foi desenvolvido na linguagem Java, utiliza a API Zxing [ZXING, 2012], que é um código aberto de leitura de códigos de barras, incluindo o QR Code. Além disso, o módulo **Mobile** utiliza a *API Mobile Payment* [PayPal, 2012], que é uma biblioteca muito simples para que vendedores possam integrar pagamentos diretamente de suas aplicações.

3. Considerações Finais

Apresentamos o sistema **Mobile Market**, que é um sistema de compras móvel rápido e seguro que utiliza as tecnologias QR Code e chaves simétricas. Uma vantagem deste sistema é o uso de criptografia de chave simétrica, que é uma técnica simples e

computacionalmente eficiente, para garantir que o QR Code lido é exatamente o mesmo QR Code gerado pela empresa que está vendendo os produtos, futuramente planeja-se testar o software com outras técnicas de criptografia, como a de chave assimétrica e com uma que é um misto entre a técnica anterior e a simétrica.

De acordo com Kieseberg et. al.(2010), é possível se prevenir ataques à processos automatizados e ataques à interação humana. Ataques à processos automatizados são: injeção SQL, injeção de comandos e fraudes. Os ataques à interação humana são: *phishing and pharming*, fraudes, leitor de software malicioso e engenharia social. Um exemplo de uma ação mal intencionada seria o de substituir ou sobrepor o QR Code original na vitrine (fraude), e com isso, redirecionar o usuário para outro site, que não o original. Todos os tipos de ataque se fiam na possibilidade de se criar QR Code com conteúdo malicioso. Na solução **Mobile Market**, todos os tipos de ataque são prevenidos via criptografia/descriptografia da informação, a não ser que o autor do ataque tenha conhecimento da chave.

Esta solução ainda está em desenvolvimento e já se efetua toda a operação de compra, no entanto, necessita de um aperfeiçoamento da interface com o usuário visando uma melhor atratividade. Para estender ainda mais as possibilidades, planeja-se futuramente, desenvolver o sistema **Mobile** para as plataformas Windows Phone e iOS.

Referências

- cert.br (2012). “Cartilha de Segurança para Internet”. <http://cartilha.cert.br/criptografia/>, Junho.
- Denso Wave Incorporate. (2012) “About QRCode.com”. <http://www.qrcode.com/en/index.html>, Outubro.
- G1 (2012). “Compras com celulares chegarão a US\$ 1 trilhão em 2017, diz pesquisa”. <http://g1.globo.com/tecnologia/noticia/2012/11/compras-com-celulares-chegarao-us-1-trilhao-em-2017-diz-pesquisa.html>, Novembro.
- Estadão. (2012) "A estratégia dos grandes: rede de supermercados inova ao colocar prateleira virtual no metrô". <http://pme.estadao.com.br/noticias/noticias,a-estrategia-dos-grandes-rede-de-supermercados-inova-ao-colocar-prateleira-virtual-no-metro,1708,0.htm>, Abril.
- Gullaksen, K. (2012). “QRGen”. <http://kenglxn.github.com/QRGen>, Outubro.
- Kieseberg, P.; Leithner, M.; Mulazzani, M.; Munroe, L.; Schrittwieser, S.; Sinha, M. e Weippl, E. (2010) “Qr Code Security”. In: Proc. of The International Workshop on Trustworthy Ubiquitous Computing (TwUC’10), Paris, France.
- Komati, K. S.; Costa, E. R. e Andrade, J. O. (2012). "Gerenciamento de Informações com QR Code e Código Hash Criptográfico". Em: Anais do XIX Simpósio de Engenharia de Produção (SIMPEP), 2012, Bauru.
- PAYPAL. (2012). “Mobile Payment Libraries”. <https://www.paypal-brasil.com.br/x/xblog/2012/02/15/mobile-payment-libraries>, Fevereiro.
- Wasserman, T. (2011) "New Security Threat: Infected QR Codes", <http://mashable.com/2011/10/20/qr-code-security-threat/>, October.
- ZXING. (2012) “ZXing - Zebra Crossing”. <http://code.google.com/p/zxing>, Outubro.

Resolução de PPLs através dos métodos Simplex e Simplex em Duas Fases

Leandro Pedrosa Rodrigues¹, Veríssimo Guimarães Júnior¹

¹Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Catalão – GO – Brasil

pedrosalpr@gmail.com, verissimoguimaraes@yahoo.com.br

Abstract. *Linear Programming Issues demand a high level of knowledge to solve them in practice. To minimize the trouble enforced by Simplex Methods of resolution it has been made a prototype in order to help understand the algorithms, showing step by step to obtain the optimal solution. The article discussed about Linear Programming and the Simplex Method of two types that are implemented in the prototype. The results for the models showed the efficiency of the tool, which helps complement and learning their methods.*

Resumo. *Problemas de Programação Linear exigem um grau de conhecimento para resolvê-los na prática. Para amenizar a dificuldade que é imposta pelos Métodos Simplex para sua resolução foi criado um protótipo no intuito de auxiliar no entendimento dos algoritmos, mostrando passo-a-passo a obtenção da solução ótima. No artigo é abordado sobre a Programação Linear e os dois tipos de Método Simplex que são aplicados no protótipo. Os resultados encontrados para os modelos mostraram a eficiência da ferramenta, o que auxiliam e complementam o aprendizado dos respectivos métodos.*

1. Introdução

A Programação Linear (PL) é um planejamento de atividades sobre funções lineares, o qual tem como objetivo obter o resultado ótimo em um determinado problema chamado de Problema de Programação Linear (PPL) [Santos 2000]. Para alcançar tal objetivo vários métodos podem ser utilizados. Dentre eles, podemos citar, o método gráfico, os métodos simplex e alguns métodos computacionais já existentes como SOLVER (Microsoft EXCEL®), AIMMS® (Paragon Decision Technology), LINDO® - Linear, INTe-ractive, and Discrete Optimizer (LINDO Systems Inc.) [Filho 2004] e o PROLIN (Universidade Federal de Viçosa - UFV).

O objetivo deste trabalho é apresentar um protótipo de uma ferramenta Web para resolver os PPLs através dos Métodos Simplex. A diferença com relação ao SOLVER e o AIMMS, além de ser gratuito, é que essa ferramenta irá mostrar passo-a-passo, cada fase da sua resolução, desde a formulação até a solução do problema. O LINDO e o PROLIN que também são gratuitos não disponibilizam esse passo-a-passo.

A importância desse material didático-pedagógico será dedicada ao ensino das disciplinas de Pesquisa Operacional e suas variações dentro dos cursos de graduação e pós-graduação da UFG (Universidade Federal de Goiás) e qualquer outra IES (Instituição de Ensino Superior) que tiver interesse no assunto abordado, visto que essa ferramenta também irá facilitar o aprendizado e o entendimento dos algoritmos Simplex com um

número maior de variáveis e restrições do problema, que dependendo desse número de variáveis, se torna inviável a explicação e o acompanhamento dentro de uma sala de aula.

2. Conceitos Básicos

Nesta seção serão apresentados os conceitos básicos para a formulação do protótipo, entre eles, Programação Linear, Método Simplex e Método Simplex em Duas Fases.

2.1. Programação Linear

Programação Linear é uma das técnicas da Pesquisa Operacional das mais utilizadas em se tratando de problemas de otimização. Os problemas de Programação Linear buscam a distribuição eficiente de recursos limitados para atender um determinado objetivo, em geral, maximizar lucros ou minimizar custos. Em se tratando de PL, esse objetivo é expresso através de uma função linear, denominada de "Função Objetivo".

2.2. Método Simplex

Segundo [Andrade 2009], são necessários 3 passos para iniciar o processo do método simplex: a formulação do problema, montagem do modelo e a solução do modelo.

A formulação do problema consiste em montar o problema de acordo com queira se produzir, se adequando pelas limitações. A montagem do modelo está na construção do modelo matemático, determinando a função objetivo com suas respectivas restrições.

A solução do modelo é a resolução do problema para se encontrar o valor ótimo. Para resolver o problema, são necessários realizar os procedimentos descritos em [Andrade 2009] para problemas de maximização.

2.3. Método Simplex em Duas Fases

Este método é utilizado quando as restrições são do tipo $(\geq)$ ou $(=)$. Para que a solução básica inicial não seja negativa, é preciso adicionar variáveis artificiais em cada uma dessas restrições, as quais não terão significado algum para o problema real, porém são necessárias para a resolução do PPL. Esta resolução utiliza o mesmo procedimento do método Simplex, mas são divididas em duas fases. A primeira fase consiste em encontrar a solução básica inicial. Após essa fase, as variáveis artificiais são eliminadas do quadro simplex dando início a segunda fase [Andrade 2009].

3. Protótipo

O protótipo foi criado a partir da linguagem PHP (Personal Home Page), que é uma linguagem de script open source de uso geral [Group 2001]. Além do PHP, foram utilizados CSS (Cascading Style Sheets), o qual é uma linguagem para estilos que define o layout de documentos HTML [HTML.net]. Em termos de armazenamento de dados não foi utilizado nenhum banco de dados sendo as informações armazenadas no servidor enquanto aplicação está sendo executada. A escolha da linguagem para a implementação do protótipo foi devida a sua disponibilidade na Web e por sua interoperabilidade da plataforma do usuário.

Para encontrar a solução de um PPL foi desenvolvido a ferramenta Web, RePPL, que tem por objetivo encontrar a solução ótima do PPL através dos Métodos Simplex e

Simplex em Duas Fases, podendo o usuário escolher passo-a-passo ou diretamente. A escolha passo-a-passo mostra desde a entrada do conjunto de equações e inequações, a introdução das variáveis de folga (restrições $\leq$) ou de excesso (restrições $\geq$) e artificiais (para o método Simplex em Duas Fases e restrições $\geq$ e $=$), a montagem do quadro simplex inicial, quais são as variáveis que devem entrar e sair da base e as transformações do quadro simplex de tal forma que as colunas das variáveis que estão na base se tornem um vetor identidade. Se a escolha for por executar diretamente, os cálculos serão realizados implicitamente e apenas será mostrado o resultado final.

O protótipo atualmente está hospedado no sítio www.leandropedrosa.com.br/repl/ e em funcionamento para utilização. Será realizado testes com os alunos que estão matriculados nos cursos de Ciência da Computação e Administração da UFG (Campus Catalão) nas disciplinas de Pesquisa Operacional e Métodos e Modelos Determinísticos em Administração. Em seguida será apresentado um questionário para medir o índice de satisfação do protótipo. As perguntas do questionário serão medidas entre 0 (zero) e 10 (dez) no aspectos de interface com o usuário e sobre a execução do protótipo.

Sobre a Interface serão avaliados: O layout (interface) da ferramenta é agradável? Os títulos, subtítulos e textos estão legíveis de serem lidos? Os títulos empregados nos menus refletem de forma objetiva o contexto em questão? Os formulários disponíveis nas telas do protótipo foram de fácil entendimento? O vocabulário empregado no protótipo lhe parece agradável ao seu entendimento? Qual o grau de satisfação que você atribui ao design da interface do protótipo?

Sobre a execução do protótipo a avaliação será: No seu ponto de vista, o protótipo manteve seu objetivo de auxiliar nos ensinamentos dos Métodos Simplex e Simplex em Duas Fases? Em relação à explicação de cada passo, foi de fácil entendimento? Você já utilizou outros métodos computacionais para a resolução de PPLs? Se já utilizou, alguns deles mostravam a resolução passo-a-passo? Em termos pedagógicos, você retornaria ao protótipo para resolver outros PPL? Você indicaria aos seus colegas, a fim de estudo, a utilização desse protótipo para a resolução de um PPL? Qual sua nota final para essa ferramenta computacional?

4. Resultado

Para demonstrar o funcionamento do protótipo na opção Simplex foi utilizado o seguinte exemplo:

$$\text{Maximizar } Z = 10x_1 + 8x_2 + 9x_3 + 7x_4$$

Sujeito a:

$$x_1 \leq 70$$

$$x_2 \leq 60$$

$$x_3 \leq 40$$

$$x_4 \leq 20$$

$$5x_1 + 4x_2 + 8x_3 + 9x_4 \leq 800$$

$$2x_1 + 6x_2 + 8x_4 \leq 200$$

$$2x_1 + 4x_2 + 2x_3 + 8x_4 \leq 12000$$

$$7x_1 + 3x_2 + 7x_4 \leq 1600$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0$$

Após a execução do protótipo RePPL, a solução obtida foi $x_1 = 70, F_2 = 50, x_3 = 40, F_4 = 20, F_5 = 90, x_2 = 10, F_7 = 11740, F_8 = 1080$ e com $Z = 1140$.

Para a opção Simplex em Duas Fases foi utilizado o seguinte exemplo:

$$\text{Minimizar } Z = 16x_1 + 12x_2 + 5x_3$$

Sujeito a:

$$8x_1 + 4x_2 + 4x_3 \geq 16$$

$$4x_1 + 6x_2 \geq 12$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$$

Executando novamente o protótipo, a solução obtida foi $x_2 = 2, x_3 = 2$ e com $Z = 34$.

5. Conclusão

Este artigo mostrou a criação de um protótipo para a resolução direta ou passo-a-passo de um PPL através dos métodos Simplex e Simplex em Duas Fases, trazendo uma alternativa para encontrar a solução ótima de modelos de otimização. Entendendo a dificuldade dos alunos em compreender os métodos e também de não encontrar outros softwares que mostrassem passo-a-passo a obtenção da solução ótima de um PPL, o protótipo RePPL conseguiu encontrar a solução ótima para os modelos apresentados. Vale salientar que essa ferramenta não substitui a abordagem e o ensinamento na sua parte escrita, mas auxilia nos seus esclarecimentos.

Com uma interface amigável, foi possível perceber que o resultado obtido proporciona um aprendizado e conhecimento sobre os métodos.

Esperando que esse trabalho traga contribuição não apenas para estudantes da UFG, mas para todos aqueles interessados no assunto abordado, fica como trabalhos futuros a implementação dos métodos Dual Simplex e Big M no protótipo RePPL.

Referências

- [Andrade 2009] Andrade, E. L. (2009). *Introdução à pesquisa operacional: métodos e modelos para análise de decisões*. LTC.
- [Filho 2004] Filho, V. J. M. F. (2004). O uso de software de modelagem aimms na solução de problemas de programação matemática. *Pesquisa Operacional*, 24(1):197–210.
- [Group 2001] Group, T. P. (2001). O que é php? url: http://php.net/manual/pt_BR/intro-whatis.php Acessado em 22 de fevereiro de 2013.
- [HTML.net] HTML.net. Css. <http://pt-br.html.net/tutorials/css/lesson1.php> Acessado em 22 de fevereiro de 2013.
- [Santos 2000] Santos, M. P. (2000). Programação linear. url: <http://www.mpsantos.com.br/plinear/arquivos/plinear.pdf> Acessado em 26 de dezembro de 2012.

Gerenciamento de melhorias de informações utilizando *Data Warehouse*

Reni Elisa da Silva Pontes¹, Fabiano Pontes Pereira da Silva², Douglas Willer
Ferrari Luz Vilela³, Gilvani Alves⁴

¹Pro-reitoria de Ensino - Instituto Federal de Educação Ciência e Tecnologia de Mato Grosso (IFMT) – Cuiabá, MT – Brasil

²Diretoria de Gestão de Tecnologia da Informação - Instituto Federal de Educação. Ciência e Tecnologia de Mato Grosso (IFMT) – Cuiabá, MT – Brasil

³Coordenação de Tecnologia da Informação - Instituto Federal de Educação. Ciência e Tecnologia de Mato Grosso (IFMT) – Campus Bela Vista - Cuiabá, MT – Brasil

⁴Coordenação de Tecnologia da Informação - Instituto Federal de Educação. Ciência e Tecnologia de Mato Grosso (IFMT) – Campus Bela Vista - Cuiabá, MT – Brasil

reni.pontes@ifmt.edu.br, fabiano.silva@ifmt.edu.br,
douglas.vilela@blv.ifmt.edu.br, gilvani.alves@blv.ifmt.edu.br

Abstract. *The data warehouse, when applied to meet the related bussiness area needs, enables organizations to integrate data from several different systems and modules and to create opportunities linked to business. This article presents a study aiming to clarify concepts and applicability related to data warehouse. It proposes requirements that contribute to the management of information improvement using the data warehouse.*

Resumo. *O Data Warehouse possibilita que as organizações integrem dados de diversos sistemas e módulos distintos, e criem oportunidades ligadas a negócios quando aplicado de maneira a atender a área relacionada. Este artigo apresenta uma pesquisa com o objetivo de esclarecer conceitos e aplicabilidade relacionados a Data Warehouse, buscando encontrar requisitos que contribuam para o gerenciamento de melhorias de informações utilizando o Data Warehouse.*

1 Introdução

Nos dias atuais a busca pelas informações que sejam precisas e que esteja alinhada a melhoria na qualidade dessas informações, faz com que busquemos maneiras para atender essa necessidade. As informações devem possuir qualidade e os resultados devem ser alinhados no sentido de integrar a necessidade de uma determinada realidade. De acordo com Favaretto (2001) a integração das funções de cada informação pode introduzir melhorias significativas.

Neste artigo a relevância do tema percebe-se pela evolução de informações que vem aumentando gradativamente e de uma maneira dinâmica na aplicabilidade de desempenho, custos das tecnologias de informação, recursos de informática disponíveis, capacidade de processamento e crescimento na facilidade de usabilidade.

Este artigo objetiva encontrar respostas baseadas em artigos para responder aos questionamentos sobre o gerenciamento de melhorias de informações utilizando *Data Warehouse*. Nesses questionamentos serão respondidas as seguintes perguntas: (1) como podemos tratar diversos dados equivalentes? (2) como fazer para padronizar as terminologias? (3) até onde ou como a informação pode ser democratizada? (4) como

agilizar o processo de decisão? (5) como aperfeiçoar tempo dos profissionais envolvidos?.

Para responder aos questionamentos serão referenciadas algumas literaturas relacionadas a *Data Warehouse*, onde possam ser identificados requisitos referentes a melhorias das informações e que possam responder aos questionamentos descritos.

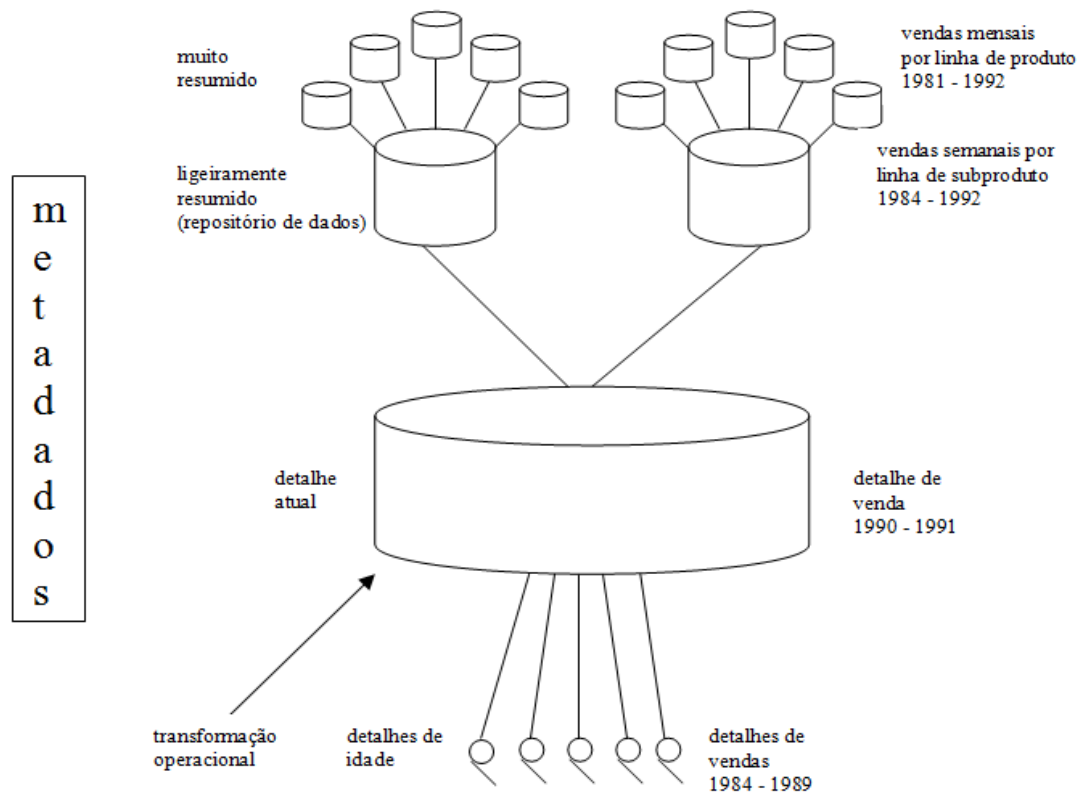
A pesquisa constitui-se de uma revisão bibliográfica através do portal capes e scholar, bem como a construção de um mapa mental e em seguida uma tabela onde possibilitará serem identificados visualmente alguns requisitos para o gerenciamento de melhorias de informações utilizando *Data Warehouse*.

2. Trabalhos Relacionados

Inmon et al. (2002) afirmam que o ambiente criado pela utilização do conceito de *Data Warehouse* permite que as empresas integrem dados de diversos sistemas e módulos distintos, e criem oportunidades de negócio a partir de sua utilização.

Nessa perspectiva, o *Data Warehouse* é um banco de dados que armazena dados utilizados diariamente nas organizações que são de potencial interesse aos tomadores de decisões. Segundo (TAURION, 1997), ao reunir informações dispersas por diversos bancos de dados e plataformas distintas, o *Data Warehouse* permite que sejam feitas análises bastante eficazes, transformando dados esparsos em informações estratégicas antes inacessíveis ou subaproveitadas. A figura 1 apresenta uma estrutura de *Data Warehouse*.

Figura 1 - Estrutura de *Data Warehouse*.



Fonte: Adaptado de (INMON, 2002)

Nos dias atuais os sistemas precisam determinar a melhor forma de extrair informações para o que está sendo chamado de "análise de dados", isto é, a capacidade de compreender e analisar as informações contidas nos sistemas de software de aplicação, conforme (LANGER, 2012). As responsabilidades dos gestores em fornecer esse tipo de informação são consideradas crescentes.

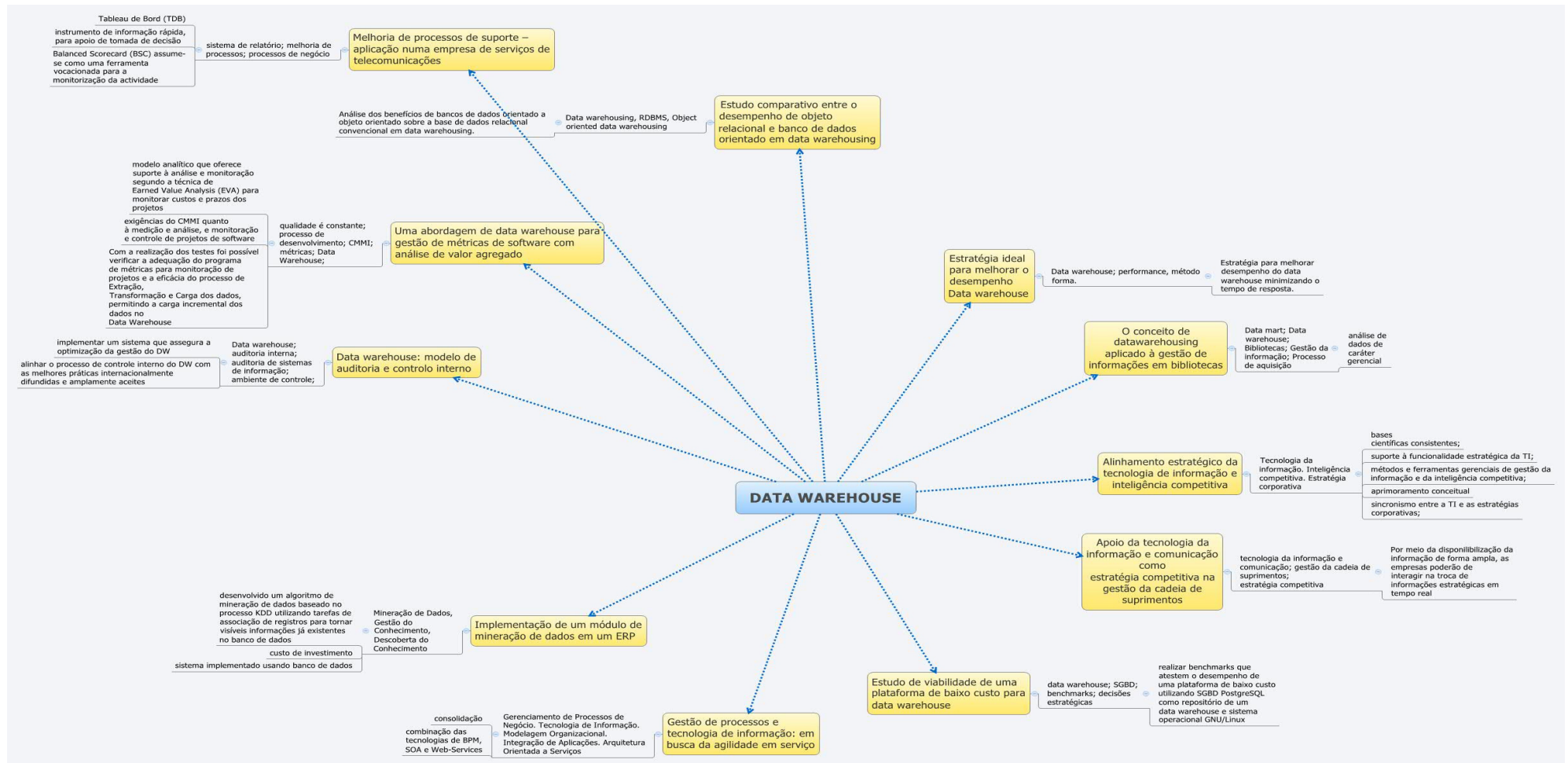
3. Metodologia

É descrito nesta seção solução encontradas para requisitos do *Data Warehouse* alinhados a melhorias das informações. A pesquisa se constituiu de análises de diversas literaturas.

3.1. Seleção das literaturas relacionadas a melhoria de *Data Warehouse*

Para a seleção das literaturas relacionadas, fez-se necessário a construção de um mapa mental, conforme figura 2, utilizando a ferramenta Xmind, onde foi possível organizar os artigos e relacioná-los ao tema proposto, denominado gerenciamento de melhorias de informações utilizando *Data Warehouse*.

Figura 2 – Visualização de literaturas relacionadas.



Nos artigos apresentados no mapa mental as melhorias encontradas foram classificadas e alinhadas em melhorias que correspondia em cada área, surgindo assim aquela que pudesse responder com afinidade as perguntas elaboradas. Cada pergunta foi representada por uma determinação ora denominada P1, P2, P3, P4, e terminando em P5, conforme abaixo:

P1 - Como podemos tratar diversos dados equivalentes?

P2 - Como fazer para padronizar as terminologias?

P3 - Até onde ou como a informação pode ser democratizada?

P4 - Como agilizar o processo de decisão?

P5 - Como aperfeiçoar tempo dos profissionais envolvidos?

Dessa forma, as perguntas estruturadas no mapa mental, bem como as respostas correspondentes, foram aplicadas consequentemente através de um contexto de instrumento avaliativo dos artigos, onde ocorreu através de palavras-chave, resumo, dentre outras informações, melhorias e resultados sobre a estrutura apresentada nos mesmos para um dado conjunto de conceitos.

3.2. Detecção e classificação das literaturas relacionadas

Após a detecção e classificação das melhorias de acordo com cada artigo pesquisado, buscou-se através de uma tabela acrescentar as perguntas aos artigos que possui a melhoria que justifique a determinada pergunta, obtendo assim as possíveis respostas, de acordo com a tabela a seguir:

Tabela 1 – Distribuição dos artigos e melhorias encontradas

Pergunta	Artigo	Melhoria encontrada	Referências
P2, P3, P4, P5	<i>Data Warehouse</i> : modelo de auditoria e controle interno	Implementar um sistema que assegura a otimização da gestão do DW; alinhar o processo de controle interno do DW com as melhores práticas internacionalmente difundidas e amplamente aceitas.	Rui Almeida Santos, 2009.
P1, P2, P3, P4, P5	Uma abordagem de <i>Data Warehouse</i> para gestão de métricas de software com análise de valor agregado	Modelo analítico que oferece suporte à análise e monitoração segundo a técnica de Earned Value Analysis (EVA) para monitorar custos e prazos dos projetos; exigências do CMMI quanto à medição e análise, e monitoração e controle de projetos de software; com a realização dos testes foi possível verificar a adequação do programa de métricas para monitoração de projetos e a eficácia do processo de extração, transformação e carga dos dados, permitindo a carga incremental dos dados no <i>Data Warehouse</i> .	Graziele Marques Mazuco, 2011.
P2, P3, P5	Melhoria de processos de suporte – aplicação numa empresa de	Tableau de Bord (TDB), instrumento de informação rápida, para apoio de tomada de decisão; Balanced Scorecard (BSC) assume-se como uma ferramenta vocacionada para a	João Ribeirinho Soares, 2011.

Pergunta	Artigo	Melhoria encontrada	Referências
	serviços de telecomunicações	monitorização da atividade.	
P1, P2, P3, P4, P5	Implementação de um módulo de mineração de dados em um ERP	Desenvolvido um algoritmo de mineração de dados baseado no processo KDD utilizando tarefas de associação de registros para tornar visíveis informações já existentes no banco de dados; custo de investimento; sistema implementado usando banco de dados.	Luiz Paulo Rech Guindani; Christiano Cadoná, 2011.
P1, P2, P3, P4, P5	Gestão de processos e tecnologia de informação: em busca da agilidade em serviço	Consolidação; combinação das tecnologias de BPM, SOA e Web-Services.	Saulo Barbará de Oliveira, Rosa Amelita Sá Menezes da Motta, Altamar Sales de Oliveira, 2012.
P2, P3, P5	Estudo de viabilidade de uma plataforma de baixo custo para <i>Data Warehouse</i>	Realizar benchmarks que atestem o desempenho de uma plataforma de baixo custo utilizando SGBD PostgreSQL como repositório de um <i>Data Warehouse</i> e sistema operacional GNU/Linux.	Eduardo Cunha de Almeida, 2004
P1, P3, P4, P5	Apoio da tecnologia da informação e comunicação como estratégia competitiva na gestão da cadeia de suprimentos	Por meio da disponibilização da informação de forma ampla, as empresas poderão interagir na troca de informações estratégicas em tempo real.	Adilson Torres ¹ ; Clóvis Luis Padoze ² ; Silvio Roberto Ignácio Pires, 2012.
P1, P2, P3, P4, P5	Alinhamento Estratégico da tecnologia de informação e inteligência competitiva	Bases científicas consistentes; suporte à funcionalidade estratégica da TI; métodos e ferramentas gerenciais de gestão da informação e da inteligência competitiva; aprimoramento conceitual; sincronismo entre a TI e as estratégias corporativas.	Leonel Cezar Rodrigues; Marcelo Jorge Fernandez, 2011.
P1, P2, P4, P5	O conceito de <i>Data Warehousing</i> aplicado à gestão de informações em bibliotecas	Análise de dados de caráter gerencial.	Maurício Ferreira Santana, 2011.

Pergunta	Artigo	Melhoria encontrada	Referências
P2, P3, P4, P5	Estudo comparativo entre o desempenho de objeto relacional e banco de dados orientado em <i>Data Warehousing</i>	Análise dos benefícios de bancos de dados orientado a objeto sobre a base de dados relacional convencional em <i>Data Warehousing</i> .	Pushpa Suri; Meenakshi Sharma, 2011.
P1, P3, P4, P5	Estratégia ideal para melhorar o desempenho <i>Data Warehouse</i>	Estratégia para melhorar desempenho do <i>Data Warehouse</i> minimizando o tempo de resposta.	Fahad Sultan; Abdul Aziz, 2010.

Baseado nas leituras dos artigos e criação da tabela percebeu-se através dessas diversas literaturas selecionadas o alinhamento relacionado com cada artigo de acordo com os questionamentos antes realizados. Cada literatura teve contribuição importante baseada em cada melhoria que cada artigo apresentou.

4. Conclusões e Trabalhos Futuros

O objetivo deste trabalho foi atingido com a obtenção de evidências que possibilitaram a elaboração de um agrupamento dos principais requisitos relacionados à *Data Warehouse*. Porém, uma revisão crítica mais extensa pode revelar requisitos e correlações sobre melhorias de informações.

O valor de *Data Warehouse* não está em colecionar dados e sim saber gerenciar os dados para serem transformados em informações úteis. No entanto, sistemas de computação como o *Data Warehouse*, permitirá maior agilidade no processo de seleção e integração das informações.

Para tanto, destacam-se ainda a quantidade de artigos pesquisados, que pode ser ampliada, e a falta de definições únicas acerca dos requisitos encontrados. Por isso, sugere-se a criação de uma ontologia que possa suprir essa lacuna, possibilitando avançar nas comparações para obter uma base de estudo mais sólida.

Como estudo futuro, sugere-se uma pesquisa específica para cada pergunta realizada neste artigo, através de ontologias.

Referências

- Almeida, E. C. (2004), “Estudo de viabilidade de uma plataforma de baixo custo para *Data Warehouse*”. Dissertação (Mestrado em Informática) – Universidade Federal do Paraná, Curitiba.
- Denzin, Normam K.; Lincoln, Yvonna S. (2006), “O Planejamento da pesquisa qualitativa. Teorias e Abordagens”. Porto Alegre: Artmed.
- Favaretto, F. (2001), “Uma contribuição ao processo de gestão da produção pelo uso da coleta automática de dados de chão de fábrica”. Tese (Doutorado em Engenharia Mecânica). Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos.
- GIL, Antonio Carlos. (2002), “Como elaborar projetos de pesquisa”. 4. ed. São Paulo: Atlas, p.175.

- Guindani, L. P. R.; Cadoná, C. (2012), “Implementação de um módulo de mineração de dados em um ERP”.
- Inmon, W., Terderman, R., Imhoff, C. (2001), “*Data Warehousing*: como transformar informações em oportunidades de negócios”. São Paulo, Brasil: Editora Berkely.
- Kimball, R. (1998), “*Data Warehouse tool kit* técnicas para a construção de *data warehouses* dimensionais”. São Paulo: Makron Books.
- Langer, A. M. (2012), “*Guide to software development*”. Springer, p. 225-259.
- Mazuco, G. M. (2011), “Uma abordagem de *Data Warehouse* para gestão de métricas de software com análise de valor agregado”. Dissertação (Graduação em Bacharelado em Ciência da Computação) – Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre.
- Oliveira, S. B. et al. (2012), “Gestão de processos e tecnologia de informação: em busca da agilidade em serviço”. *Revista eletrônica de gestão organizacional*, vol. 10, n. 1, p. 172-194.
- Rodrigues, L. C.; Fernandez, M. J. (2011), “Alinhamento estratégico da tecnologia de informação e inteligência competitiva”. *Revista inteligência competitiva*, vol. 1, n. 3, p. 328-344.
- Santos, R. A. (2009), “*Data Warehouse*: modelo de auditoria e controlo interno”. Dissertação (Mestrado em Sistemas Integrados de Apoio à Decisão (Business Intelligence) – Instituto Universitário de Lisboa.
- Santana, M. F. (2011), “O conceito de *Data Warehousing* aplicado à gestão de informações em bibliotecas”. *Revista digital de biblioteconomia e ciência da informação*, vol. 8, n. 2, p. 131-134.
- Soares, J. R. (2011), “Melhoria de processos de suporte – aplicação numa empresa de serviços de telecomunicações”. Dissertação (Mestrado em Engenharia Mecânica) – Faculdade de Engenharia da Universidade do Porto.
- Suri, P.; Sharma, M. M. (2011), “*A comparative study between the performance of relational & object oriented database in Data Warehousing*”. *International Journal of Database Management Systems*, vol. 3, n. 2.
- Sultan, F.; Aziz, A. (2010), “*Ideal strategy to improve data warehouse performance. International journal on computer science and engineering*”, vol. 2, n. 2, p. 409-415.
- Taurion, C. (1998), “*Data Warehouse*: vale a pena gastar milhões investindo em um?” *Developers’ Magazine*, ano 2, n. 18.
- Torres, A. et al. (2012), “Apoio da tecnologia da informação e comunicação como estratégia competitiva na gestão da cadeia de suprimentos”. *Revista Gestão Industrial*, vol. 8, n. 1, p. 22-37.

Application programming interface Library para Algoritmo Genético com Operador Transgênico

**Fabício Alves Rodrigues², Stéfanne Alves de Souza¹, Renan Vinícius Aranha¹
Ana Paula Freitas Vilela Boaventura¹**

¹Curso de Ciência da Computação - Universidade Federal de Goiás/Jataí (UFG)
BR364 Km 192, Setor Industrial, Jataí - GO - Brasil

²Pós-Graduação em Ciência da Computação - Universidade Federal de São Carlos (UFSCar)
Rodovia Washington Luís, km 235 - SP-310 São Carlos - São Paulo - Brasil

fabricao1989@gmail.com, tefi70@hotmail.com

Abstract. *This paper presents features and operating mode API (Application programming interface) of the genetic algorithm (GA) of Holland in 1975, which is based on the natural selection process proposed by Charles Darwin with the aim of get results favorable to complex problems. The API Also addresses a new genetic operator proposed in 2011 for Amaral and Hruschka, which is a genetic algorithm with a transgenic operator. Such change allows the population be diversified, and have got faster and favorable results to the problem. The API has been tested to mining data from a dataset from car traffic. The result was a mining with a textit fitness 100 % and the API has correctly executed the steps described by AG.*

Resumo. *Este artigo apresenta características e o modo de funcionamento da API(Application programming interface) do algoritmo genético (AG) de Holland em 1975, que baseia-se no processo de seleção natural proposto por Charles Darwin com o intuito de se obter resultados favoráveis para problemas complexos. A API Ainda aborda a proposta feita em 2011 por Amaral e Hruschka, que propuseram um algoritmo genético com um operador transgênico. Tal alteração permite que a população seja diversificada, além de se obter mais rápido resultado favorável ao problema. A API foi testada para minerar dados de um dataset referente a tráfego de carros. O resultado foi uma mineração com um fitness de 100% e a API executou corretamente os passos descritos pelo AG.*

1. Introdução

Nos últimos anos, as heurísticas têm tido destaque como uma das abordagens mais promissoras para solução de problemas de programação inteira e combinatórios. Dentre as técnicas heurísticas mais sofisticadas, existem as chamadas meta-heurísticas, com varias aplicações nas diferentes áreas da ciência [Rodrigues et al. 2004]. As principais meta-heurísticas que apresentam grande potencial para encontrar soluções de problemas são: Busca Tabu (BT) [Glover 1977], *Ant Colony Optimization* (ACO) [Dorigo et al. 1996], *GRASP (Greedy Randomized Adaptive Search Procedure)* [Feo and Resende 1989], *Particle Swarm Optimization* (PSO) [Kennedy and Eberhart 1995] e Algoritmo Genético (AG) [Holland 1975].

AG tem sido muito usado para resolver problemas combinatórios e, em 2011, os autores [Amaral and Hruschka 2011] propuseram um algoritmo genético com um operador transgênico. O presente trabalho tem como objetivo descrever uma *Application programming interface Library* (API) do Algoritmo Genético com operador transgênico para validar a API será feita uma mineração de dados. O presente trabalho tem como objetivo descrever um *Application programming interface Library* (API) do Algoritmo Genético com operador transgênico. Para validar a API, será demonstrado um exemplo utilizando-se a mineração de dados. Na seção 2, será descrito o funcionamento do algoritmo genético e do operador transgênico, enquanto a seção 3 conterá a descrição da API do AG com operador transgênico. Por último, serão apresentados resultados e conclusões.

2. Algoritmo Genético e Operador Transgênico

Os Algoritmos Genéticos são técnicas de busca que utilizam o processo da seleção natural proposto por Charles Darwin. Os AGs inicialmente foram descritos por John Holland [Holland 1975] e são bastante utilizados em problemas combinatórios.

Os AGs utilizam operadores para fazer sua busca, os principais operadores são seleção de indivíduos para o *crossover*, o mais simples é o método da roleta no qual cada indivíduo da população recebe uma fatia da roleta proporcional ao (*Fitness*) do indivíduo. Em seguida é feito o sorteio aleatório desses indivíduos e os selecionados fazem parte do operador de *crossover* [Man et al. 2001].

Crossover é a troca de genes entre os indivíduos selecionados. Um ponto de cruzamento é escolhido aleatoriamente e a partir dele é feita a troca genética dos indivíduos. As informações genéticas anteriores a este ponto de um dos pais é ligado as informações posteriores do outro pai gerando a cada cruzamento dois novos filhos, que são colocados na população auxiliar [Rezende 2005]. O operador mutação é necessário para que a população seja diversificada, permite chegar mais rápido ao objetivo da busca e para fugir dos mínimos locais, alterando aleatoriamente um ou mais genes dos indivíduos gerados pelo cruzamento.

Neste trabalho será usando a estratégia mais equilibrada para reinserção, essa estratégia parece melhorar o desempenho do AG, ela é chamada de melhores pais e filhos onde são escolhidos $N/2$ melhores indivíduos geradores (pais) e $N/2$ melhores indivíduos criados (Filhos) no qual são inseridos na população gerando uma nova geração [Man et al. 2001].

Em 2011 foi proposto por Laurence [Amaral and Hruschka 2011] um algoritmo genético com um operador transgênico, esse novo operador para o algoritmo genético foi inspirado pela engenharia genética, onde permite manipular o material genético de indivíduos adicionando características que acredita-se ser a mais importante. Essa mesma atividade que é feita por engenheiros genéticos foi introduzido no AG, assim identificasse qual é o gene ou genes mais importante então esses genes dos N melhores indivíduos são passados para M outros indivíduos da população.

3. *Application programming interface Library Transgenic Genetic Algorithm*

A API Java proposta neste artigo segue os princípios descritos por Holland em seu trabalho e também contém o operador transgênico descritos por Amaral e Hruschka.

A API apresenta duas classes (*Population* e *GeneticAlgorithm*), duas interfaces (*GeneImpl* e *IndividualImpl*). A classe *Population* contém sete métodos, são eles: *generateIndividuals()* que gera os indivíduos na população, *greaterIndividual()* que retorna o indivíduo com maior *fitness*, *smallestIndividual()* que retorna o indivíduo com menor *fitness*, *getIndividual(int pos)* retorna o indivíduo da população que está na posição requerida, *getIndividuals()* retorna todos os indivíduos da população, *setIndividual(IndividualImpl ind, int pos)* modifica o indivíduo na posição requerida pelo indivíduo passado por parâmetro e *setIndividuals(IndividualImpl ind[])* modifica toda a população pelo vetor de indivíduos passado no parâmetro. A classe *Population* será utilizada pela outra classe chamada *GeneticAlgorithm*.

A classe *GeneticAlgorithm* contém os operadores descritos na seção 2, também apresenta o método *bestIndividual()* que retorna o melhor resultado encontrado pelo Algoritmo Genético, para utilizar essa classe requer que o usuário passe tamanho da população, um exemplo de indivíduo implementado, já que o mesmo é somente uma *interface* Java, também é necessário passar uma *long* que corresponde a *Seed*, caso o problema necessite executar mais de uma vez usando a mesma população inicial e um valor booleano onde (*True*: se o problema é de maximização e *False* se for de minimização).

O gene e o indivíduo da API são representados por *interface*, pois os genes e indivíduos são diferentes para cada problema que o usuário pretenda resolver. A *interface GeneImpl* contém os seguintes métodos que os usuários terão que implementar, *makeMutation()* onde deve ser implementado como será feita a mutação no gene, *createGene(Random seed)* ou *createGene()* método no qual será gerado o valor aleatório do gene, *G getValue()* retorna o valor do gene (*G* representa o tipo genérico da variável valor deve ser dito na implementação do gene, exemplo *public GeneImpl<Integer>*) e *setValue(G valor)* que deve modificar o valor do gene.

A *interface IndividualImpl* contém os seguintes métodos que os usuários terão que implementar *createGenes(Random seed)* esse método deve gerar os *N* genes que representa o indivíduo do problema a ser resolvido, *getGenes()* deve retornar todos os genes que representa o indivíduo, *setGenes(GeneImpl genes[])* deverá modificar todos os genes do indivíduo pelo vetor que será recebido por parâmetro, *getGene(int pos)* deverá retornar o gene que está na posição requerida, *setGene(int pos, GeneImpl g)* deverá modificar o gene na posição requerida pelo gene 'g' passado pelo parâmetro, *getFitness()* deverá calcular o valor do *fitness* e retorna-lo e *Mutation(int pos)* que deverá requisitar a classe *Gene*, que deve ser implementada pelo usuário, a execução de uma mutação no gene que é representado na posição 'pos'.

4. Resultados e Conclusões

Para validar a API, foi feita uma mineração de dados em um *dataset* que contém dados de trânsito, que pode ser encontrado na *abaTools and Files* do link www.fabricioarodrigues.wix.com/researches. Neste *dataset* requer encontrar uma regra *IF-THEN* que melhor represente se houve ou não congestionamento com base nos dados do *dataset*. A avaliação (*Fitness*) do AG foi formulado como sendo $TP * 0.5 + FP * 0.5$, onde *TP* são registros classificados corretamente e *FP* são registros que foram ditos como não pertencentes à classe avaliada e realmente não pertenciam. O AG foi executado com taxa de cruzamento de 100% e taxa de mutação de 2%. Eram

sorteados 2 indivíduos para cruzamento e foi utilizado a reinserção balanceada, também conhecida como melhores pais e filhos. Os genes e cromossomo usados neste artigo são os mesmo do trabalho [Cunha et al. 2012].

A API executou corretamente cada passo descrito por Holland e com o operador transgênico descrito por [Amaral and Hruschka 2011]. Apresentou ainda um resultado satisfatório, com resultados de *fitness* de 100% foi executadas 180 gerações o AG foi executado 100 vezes e as regras encontradas foram *IF Fluxo >= 51 THEN Congestionado* e *IF Fluxo < 50 THEN Não Congestionado*. Pode se concluir que, com o uso de uma API do AG transgênico, como o descrito nesse trabalho, pode-se obter resultados satisfatórios que são capazes de ajudar pesquisadores a encontrar resultados para problemas combinatórios. Como trabalhos futuros iremos implementar outras funcionalidades de seleção e cruzamento, pois usamos aqui somente uma das estratégias para os mesmos. Ainda, será feita a documentação apropriada para a API descrita neste artigo.

Referências

- Al-Hinai, N. and ElMekkawy, T. (2011). Robust and stable flexible job shop scheduling with random machine breakdowns using a hybrid genetic algorithm. *International Journal of Production Economics*, 132(2):279–291.
- Amaral, L. and Hruschka, E. (2011). Transgenic, an operator for evolutionary algorithms. In *Evolutionary Computation (CEC), 2011 IEEE Congress on*, pages 1308–1314.
- Cunha, L. C., Rodrigues, F. A., and de Souza Alencar, W. (2012). Semáforo inteligente com lógica fuzzy baseado na mineração de dados por algoritmo genético transgênico. In *EATI 2012*.
- Dorigo, M., Maniezzo, V., and Colorni, A. (1996). Ant system: optimization by a colony of cooperating agents. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 26(1):29–41.
- Feo, T. and Resende, M. (1989). A probabilistic heuristic for a computationally difficult set covering problem. *Operations research letters*, 8(2):67–71.
- Glover, F. (1977). Heuristics for integer programming using surrogate constraints. *Decision Sciences*, 8(1):156–166.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, MI, USA.
- Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. In *Neural Networks, 1995. Proceedings., IEEE International Conference on*, volume 4, pages 1942–1948. IEEE.
- Man, K. F., Tang, K. S., and Kwong, S. (2001). *Genetic Algorithms: Concepts and Designs*. Springer, 3 edition.
- Rezende, S. O. (2005). *Sistemas inteligentes: fundamentos e aplicações*. Manole Ltda, Barueri, SP.
- Rodrigues, F., Leite, H., Santos, H., de Souza, A., and Silva, G. (2004). Metaheurística algoritmo genético para solução de problemas de planejamento florestal com restrições de integridade. *Revista Árvore*, 28(2):233–245.