

2014 enacomp

XI Encontro Anual de Computação

Sistemas embarcados:
Novas visões de desenvolvimento

Anais

Apoio:

FAPEG
FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE GOIÁS

GOVERNO DE GOIÁS
Fazendo o melhor pra você.

CAPES

Ministério da Educação

GOVERNO FEDERAL
BRASIL
PAÍS RICO É PAÍS SEM POBREZA

SBC

CECCAC
COORDENAÇÃO DE EXTENSÃO E CULTURA DO CÂMPUS CATALÃO

UFG

Realização:

UFG
REGIONAL CATALÃO
Universidade Federal de Goiás

Catalão - Goiás
Av. Dr. Lamartine P. Avelar, 1120
Setor Universitário - CEP 75704-020
http://www.enacomp.com.br

DCC

Anais do XI Encontro Anual de Computação

Editores:
Tércio Alberto dos Santos Filho
Universidade Federal de Goiás – UFG

Sérgio Francisco da Silva
Universidade Federal de Goiás – UFG

Capa:
Edmar da Silva Purcina

Catalão – GO,
2014

Comissão Organizadora

Coordenador:

- Tércio Alberto dos Santos Filho

Vice-Coordenador:

- Sérgio Francisco da Silva

Alunos:

- Anelisa Pereira da Silva
- Carlos Heitor Sanches
- Charlliston Adrianni
- Diego Martins de Almeida
- Edmar da Silva Purcina
- Erick Dutra
- Felipe Otero
- Jeferson Rech Brunetta
- Jéssica da Silva Guimarães
- João Augusto Júnior
- Lara Rosa de Lima
- Letícia Alcântara Lima
- Lorena Aparecida Rezende
- Marcos Vinícios Tomas de Oliveira
- Maicon de Jesus Lima
- Melque Henrique
- Paulo Henrique da Silva Azevedo
- Rafael Rosa da Fonseca

Comitê Científico

- Antônio Carlos de Oliveira Júnior (*Chair*) – CAC/UFG
- César Augusto Lins de Oliveira – UFPE
- Crescencio Rodrigues Lima Neto – IFB
- Dalton Matsuo Tavares – CAC/UFG
- Fabrício da Costa – FACISA
- Felipe Domingos – UFMG
- Gleibson Rodrigo Silva de Oliveira – UFPE
- Henrique Emanuel Mostaert Rebêlo – UFPE
- Iuri Santos Souza – UFBA
- Iwens Gervasio Sene Junior – UFG
- José Jorge Lima Dias Júnior – UFPB
- Jeneffer Cristine Ferreira – UFPE
- Juliana de Albuquerque Gonçalves Saraiva – UAG
- Liliane Nascimento Vale – CAC/UFG
- Luanna Lopes Lobato – CAC/UFG
- Maicon Aparecido Sartin – UNEMAT
- Márcio Antônio Duarte – CAC/UFG
- Marcos Antonio Estremote – UNESP-FEIS
- Marcos Aurélio Batista – CAC/UFG
- Muriel de Souza Godoi – UTFPR
- Nádia Félix Felipe da Silva – ICMC/USP
- Rafael de Amorim Silva – UFPE
- Renata Maria de Souza Santos – FACOL
- Renato de Freitas Bulcao Neto – UFG
- Rodrigo Rocha Gomes Souza – UFBA
- Thiago Jabur Bittar – CAC/UFG
- Thiago Souto Mendes – IFBA
- Tiago da Silva Almeida – UNESP/FEIS

- Vaston Gonçalves da Costa – CAC/UFG
- Vinícius da Cunha Martins Borges – UFG
- Wesley Barbosa Thereza – UNEMAT
- Wylliams Barbosa Santos – UFPE

Apresentação

O Encontro Anual de Computação (ENACOMP) é um evento local voltado principalmente para estudantes em fases de iniciação científica e tecnológica. O ENACOMP abrange praticamente todos os tópicos de pesquisa em Ciência da Computação envolvendo teórias, métodos, experimentações e desenvolvimentos tecnológicos. A cada ano o evento tem uma temática específica com o intuíto único de direcionar as palestras e mini-cursos para uma área de pesquisa de grande destaque atual.

Neste ano o tema do evento foi: "Sistemas Embarcados: novas visões de desenvolvimento". As palestras e mini-cursos desta girar-se-ão em torno de métodos computacionais para desenvolvimento em sistemas embarcados, ferramentas de síntese de circuitos digitais, redes de comunição, redes do futuro (tais como, redes de sensores) e desenvolvimento de aplicações.

Este livro contém os 14 artigos aceitos para apresentação no ENACOMP 2014 em sua versão completa, os quais tratam de vários temas de pesquisa e desenvolvimento em Ciência da Computação. Os artigos também estão disponibilizados eletronicamente no sítio: `http:www.enacomp.com.br`.

Coordenação do XI ENACOMP:
Tércio Alberto dos Santos Filho (coordenador)
Sérgio Francisco da Silva (vice-coordenador)

Agradecimento

Agradecemos à CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) e à FAPEG (Fundação de Amparo à Pesquisa do Estado de Goiás) pelo financiamento do ENACOMP 2014.

Sumário

Uma Comparação Teórica entre dois Modelos de Programação Paralela em GPU

Lauro Cássio Martins de Paula[1]

[1] Instituto de Ciências Matemáticas e de Computação – Universidade de São Paulo (USP)
CEP: 13.566-590 – São Carlos – SP – Brazil

laurocassio21@gmail.com

Abstract. *This paper presents a comparison between two parallel architectures: Compute Unified Device Architecture (CUDA) e Open Computing Language (OpenCL). Some works in the literature have presented a computational performance comparison of the two architectures. However, there is not some complete and recent paper that highlights clearly which architecture can be considered the most efficient. Thus, the goal is to make a comparison only in level of* $hardware$, $software$, *technological trends and ease of use, highlighting one that may present the best performance in general. To this end, we describe the main works that have used at least one of the architectures. It was observed that the choice of OpenCL may seem more obvious for being a heterogeneous system. Nevertheless, it was concluded that CUDA, although it can be used only in graphics cards from* $NVIDIA^{®}$, *has been a reference and more used recently.*

Resumo. *Apresenta-se neste trabalho uma comparação entre dois modelos utilizados para programação paralela: Compute Unified Device Architecture (CUDA) e Open Computing Language (OpenCL). Alguns trabalhos na literatura apresentaram uma comparação de desempenho computacional entre esses dois modelos. Entretanto, ainda não existe algum artigo recente e completo que destaca claramente qual modelo, de fato, pode ser considerado o mais eficiente. Portanto, o objetivo deste trabalho é realizar uma comparação apenas em nível de* $hardware$, $software$, *tendências tecnológicas e facilidades de utilização, evidenciando aquele que pode apresentar o melhor desempenho de uma maneira geral. Para tal, descreve-se os principais trabalhos que já fizeram uso de pelo menos um dos modelos. Observou-se que, por ser um sistema heterogêneo, a escolha do OpenCL pode parecer mais óbvia. No entanto, foi possível concluir que CUDA, apesar de poder ser utilizado apenas nas placas gráficas da* $NVIDIA^{®}$, *tem sido uma referência e mais utilizado ultimamente.*

1. Introdução

A Computação Paralela (CP) tem contribuído bastante para diversas áreas da ciência, que vão desde simulações computacionais a aplicações científicas. A CP é uma forma eficiente do processamento da informação, com ênfase na exploração de eventos concorrentes no processo computacional [Smith 2011]. Com o avanço da tecnologia, novas arquiteturas computacionais têm sido desenvolvidas. Soluções com vários processadores em uma mesma placa vêm sendo elaboradas, e processadores com vários núcleos de processamento são a nova tendência tecnológica na atualidade [$CUDA^{TM}$ 2013].

Atualmente têm-se desenvolvido dois tipos de processadores: $Multi - core$ e $Many - core$. Os processadores $Multi - core$, normalmente, contém poucos núcleos, porém com um grande poder de processamento [Stallings 2002]. Tais processadores visam minimizar a latência de memória, reservando uma parte do chip para memória $cache$, e permitem um uso moderado de linhas de execução ($threads$) [Smith 2011]. Os $Many - core$ são desenvolvidos com dezenas ou centenas de núcleos mais simples, otimizados para uma maior vazão na execução de instruções, executando centenas ou até mesmo milhares de $threads$ [Kirk and Hwu 2011]. As *Graphics Processing Units* (GPU) são exemplo de processadores $Many - core$ [Paula 2013].

De forma correspondente à evolução do $hardware$, novos modelos de programação paralela têm sido elaborados, destacando-se dois deles: *Compute Unified Device Architecture* (CUDA) [$CUDA^{TM}$ 2013] e *Open Computing Language* (OpenCL) [Tsuchiyama et al. 2010]. Tais modelos, devido a uma ampla disponibilidade de *Application Programming Interfaces* (API), permitem que aplicações possam ser executadas mais facilmente na GPU [Gaioso et al. 2013].

Muitos trabalhos na literatura têm utilizado CUDA e (ou) OpenCL para a solução de diversos tipos de problemas paralelizáveis. Entretanto, uma comparação teórica e tecnológica entre ambos os modelos ainda tem sido pouco investigada. Após uma extensa pesquisa bibliográfica, foi possível observar que existem trabalhos que realizam apenas comparações de desempenho computacional entre ambos. Alguns mostram que CUDA supera o OpenCL para a maioria das aplicações [Karimi et al. 2010]. Por outro lado, outros evidenciam que o OpenCL pode ser uma boa alternativa em relação a CUDA [Fang et al. 2011]. Portanto, o objetivo deste trabalho é realizar uma comparação entre CUDA e OpenCL apenas em relação à aspectos de $hardware$, $software$, tendências tecnológicas e facilidades de utilização. Foi possível concluir que CUDA, embora seja uma tecnologia proprietária, tem sido uma referência e pode ser considerada mais eficiente.

Este artigo está organizado da seguinte forma. A Seção 2 descreve os principais detalhes sobre uma unidade de processamento gráfico (GPU). Os principais aspectos sobre CUDA são abordados na Seção 3. A Seção 4 detalha o OpenCL. Por fim, a Seção 5 mostra as conclusões do trabalho.

2. Unidade de Processamento Gráfico

As GPUs foram inicialmente desenvolvidas como uma tecnologia orientada à vazão, otimizada para cálculos de uso itensivo de dados, onde muitas operações idênticas podem ser realizadas em paralelo sobre diferentes dados (*Single Instruction Multiple Data - SIMD*) [Paula et al. 2013a]. Diferente de uma CPU $multicore$, a qual executa algumas $threads$ em paralelo, a GPU foi projetada para executar milhares de $threads$ [Paula et al. 2014].

Por um lado, as GPUs são melhores adaptadas para endereçar problemas que podem ser expressos através de cálculos realizados de forma paralela [Smith 2011]. Como o mesmo programa é executado para cada elemento de dado, há menos requisitos referentes a controles de fluxo e, exatamente por ser executado em muitos elementos de dados, a latência de acesso à memória pode ser ocultada pela realização de cálculos. Além do desempenho, as GPUs contam com um importante fator para seu sucesso: a presença de

mercado. Ter forte presença de mercado é fundamental para o sucesso de uma arquitetura paralela [Kirk and Hwu 2011].

Por outro lado, como toda tecnologia, as GPUs possuem suas limitações. Dependendo do volume de dados, o desempenho computacional da GPU pode se mostrar inferior quando comparado ao desempenho da CPU [Paula et al. 2013a]. Isso implica que a quantidade de dados a serem transferidos para a memória da GPU deve ser levado em consideração, devido à existência de um $overhead$ associado à paralelização das tarefas na GPU [$CUDA^{TM}$ 2009a], [Gaioso et al. 2013]. Fatores em relação ao tempo de acesso em memória também podem influenciar no desempenho computacional. Ou seja, o acesso à memória global da GPU geralmente apresenta uma alta latência e pode estar sujeito a um acesso aglutinado aos dados em memória [$CUDA^{TM}$ 2009a].

A $NVIDIA^{®}$ e a $AMD^{®}$ são exemplos de empresas que desenvolvem GPUs e disputam o mercado de computação paralela. Como mostra as Seções 3 e 4, linguagens de programação específicas para a GPU foram desenvolvidas por essas duas empresas.

3. Arquitetura para Dispositivos de Computação Unificada

CUDA foi a primeira API, criada pela $NVIDIA^{®}$ em 2006, a permitir que a GPU pudesse ser utilizada para uma ampla variedade de aplicações [$CUDA^{TM}$ 2013]. CUDA é suportada por todas as placas gráficas da $NVIDIA^{®}$, que são extremamente paralelas, possuindo muitos núcleos com diversas memórias $cache$ e uma memória compartilhada por todos os núcleos [$CUDA^{TM}$ 2009a]. No ambiente de programação CUDA, o sistema computacional distingue entre o que é executado na CPU ($host$) e o que é executado na GPU ($device$). Um programa em CUDA consiste em partes executadas no $host$ e outras partes executadas no $device$. A separação fica a cargo do compilador da $NVIDIA^{®}$ ($nvcc$) durante a compilação. O código em CUDA é uma extensão da linguagem computacional C (CUDA-C), onde algumas palavras-chave são utilizadas para rotular as funções paralelas ($kernels$) e suas estruturas de dados [Kirk and Hwu 2011].

A implementação de uma função a ser executada em paralelo pelas $threads$ nos núcleos da GPU é chamada $kernel$. Os $kernels$ normalmente geram um grande número de $threads$ para explorar o paralelismo de dados. O número de $threads$ é especificado pelo programador na chamada da função. Quando um $kernel$ é disparado, ele é executado como uma grade ($grid$) de $threads$ paralelas [$CUDA^{TM}$ 2013]. Como ilustra a Figura 1, as $threads$ em um $grid$ são organizadas em uma hierarquia de dois níveis, onde cada $grid$ consiste em um ou mais blocos de $threads$.

Desde o seu surgimento, alguns trabalhos têm utilizado CUDA para a paralelização de vários tipos de problemas. Por exemplo, Paula *et al.* (2013b) utilizaram CUDA-C para paralelizar o método BiCGStab(2), utilizado para solução de sistemas lineares. Paula *et al.* (2013a) propuseram uma estratégia de paralelização para a fase 2 do Algoritmo das Projeções Sucessivas utilizando CUDA-C. Paula (2013) utilizou a estratégia proposta em [Paula et al. 2013b] e apresentou uma comparação entre métodos iterativos na solução de sistemas lineares grandes e esparsos. Por fim, Gaioso *et al.* (2013), utilizando CUDA-C, apresentaram uma paralelização para o algoritmo Floyd-Warshall, utilizado para encontrar os caminhos mínimos entre todos os pares de vértices em um grafo.

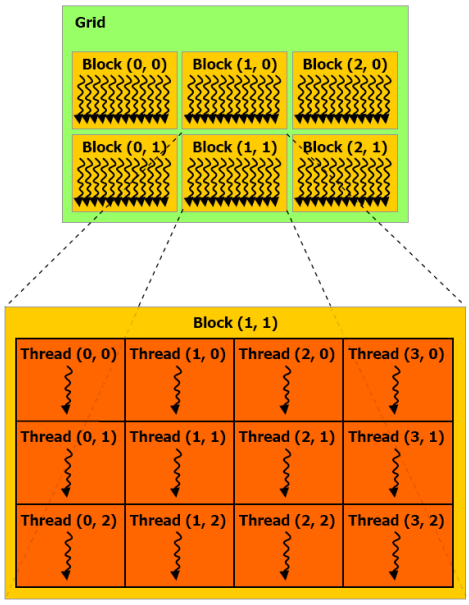


Figura 1. $Grid$ **com vários blocos de** $threads$ **[**$CUDA^{TM}$ **2013].**

Recentemente, a $MathWorks^{®}$ [Little and Moler 2013] desenvolveu um plugin capaz de fazer a integração entre CUDA e MATLAB. Fazer uso do MATLAB para computação em GPU pode permitir que aplicações sejam aceleradas mais facilmente [Little and Moler 2013]. As GPUs podem ser utilizadas com MATLAB por meio do *Parallel Computing Toolbox* (PCT). O PCT fornece uma maneira eficiente para acelerar códigos na linguagem MATLAB, executando-os em uma GPU. Para isso, o programador deve alterar o tipo de dado para entrada de uma função para utilizar os comandos (funções) do MATLAB que foram sobracarregados ($GPUArray$). Por meio da função $GPUArray$ é possível alocar dados na memória da GPU e fazer chamadas a várias funções do MATLAB, que são executadas nos núcleos de processamento da GPU. Além disso, os desenvolvedores podem fazer uso da interface $CUDAKernel$ no PCT para integrar seus códigos em CUDA-C com o MATLAB [Reese and Zaranek 2011]. O desenvolvimento de aplicações a serem executadas na GPU utilizando o PCT é, geralmente, mais fácil e rápido do que utilizar a linguagem CUDA-C [Liu et al. 2013]. Isso ocorre porque os aspectos de exploração de paralelismo são realizados pelo próprio PCT [Little and Moler 2013]. Entretanto, a organização e o número de $threads$ a serem executadas nos núcleos da GPU não podem ser gerenciados manualmente pelo programador. Ainda, é importante ressaltar que, para poder ser utilizado, o PCT requer uma placa gráfica da $NVIDIA^{®}$.

Após a integração CUDA-MATLAB, alguns trabalhos têm utilizado essa tecnologia. Por exemplo, a $NVIDIA^{®}$ (2007) lançou um livro que demonstra como programas desenvolvidos em MATLAB podem ser acelerados usando suas GPUs [$CUDA^{TM}$ 2007]. Simek e Asn (2008) apresentaram uma implementação em MATLAB com CUDA para compressão de imagens médicas [Simek and Asn 2008]. Mais recentemente, Liu *et al.* [Liu et al. 2013] apresentaram uma pesquisa e comparação de programação usando GPUs em MATLAB.

Com base nesses resultados, nota-se que, futuramente, o PCT poderá ser mais utilizado devido ao fato de permitir que um código na linguagem MATLAB possa ser mais facilmente paralelizado. Logo, ao invés de implementar uma função $kernel$ e definir a quantidade e a organização de $threads$ em blocos, o programador deve apenas identificar quais partes de seu código são paralelizáveis e fazer uso das funções padrão do MATLAB.

4. Linguagem de Computação Aberta

OpenCL é um padrão aberto, mantido pelo *Khronos Group*®, que permite o uso de GPUs para desenvolvimento de aplicações paralelas. Ele também permite que os desenvolvedores escrevam códigos de programação heterogêneos, fazendo com que estes programas consigam aproveitar tanto os recursos de processamento das CPUs quanto das GPUs. Além disso, permite programação paralela usando paralelismo de dados e de tarefas [Tsuchiyama et al. 2010].

Apesar de se tratar de um sistema aberto, o *Khronos Group*® é responsável pela padronização de alguns parâmetros. O *Khronos Group*® anunciou recentemente uma versão atualizada do OpenCL: O OpenCL 2.0, que é a mais recente evolução do padrão OpenCL, projetado para simplificar ainda mais a programação multi-plataforma, permitindo uma variedade de algoritmos e padrões de programação para serem facilmente acelerados [Khronos 2013]. Os dispositivos em OpenCL podem ou não compartilhar memória com a CPU e, normalmente, têm um conjunto de instruções de máquina diferente [Stone et al. 2010]. As APIs fornecidas pelo OpenCL incluem funções para enumerar os dispositivos disponíveis (CPU, GPU e outros aceleradores), gerenciar as alocações de memória, realizar transferências de dados entre CPU e GPU, disparar $kernels$, para serem executados nos núcleos da GPU, e verificar erros.

Recentemente, o OpenCL também têm oferecido suporte para CUDA, permitindo o desenvolvimento de aplicativos para serem executados em diversas plataformas [$CUDA^{TM}$ 2009b]. Ainda, por meio de suas diversas APIs, os desenvolvedores podem fazer chamadas à funções $kernels$ utilizando um subconjunto limitado da linguagem de programação C. Um programa em OpenCL consiste em $kernels$, que são executados pelo(s) $device(s)$, e $host$, que gerencia a execução dos $kernels$. Os $kernels$ são executados por $workitens$. Os $workitens$ são agrupados em $workgroups$. Os $workgroups$ são organizados em um $NDRange$. A Figura 2 mostra a organização dos $workitens$ e $workgroups$ em um $NDRange$.

Após seu surgimento, alguns trabalhos têm utilizado o OpenCL na tentativa de aumentar o desempenho computacional de problemas paralelizáveis. Entre eles, pode-se citar o trabalho de Komatsu *et al.* (2010), que apresentaram uma avaliação de desempenho e portabilidade de programas em OpenCL [Komatsu et al. 2010]. Barak *et al.* (2010) apresentaram algumas aplicações em OpenCL executadas em $clusters$ com muitas GPUs [Barak et al. 2010]. Mais recentemente, Suganuma *et al.* (2013) descreveram suas experiências em programação OpenCL para obter um desempenho escalável para um ambiente de computação heterogênea e distribuída [Suganuma et al. 2013].

Enquanto CUDA é mantido e aprimorado apenas pela NVIDIA®, o OpenCL é suportado por fabricantes como, por exemplo: AMD®, $NVIDIA$®, $APPLE$®, $INTEL$® e IBM®. No entanto, apesar de se tratar de um modelo heterogêneo, permitindo o gerenciamento para portabilidade em multiplataformas, o OpenCL pode se mos-

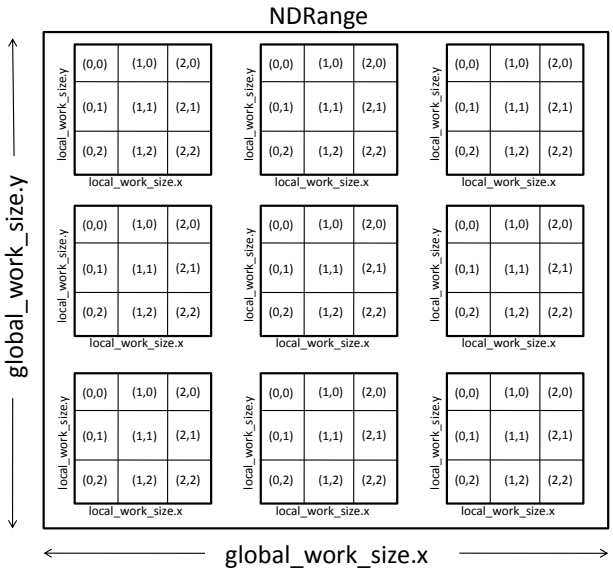


Figura 2. Organização dos $workgroups$ **e** $workitens$ **em um** $NDRange$**.**

trar um tanto quanto complexo em comparação a CUDA. Além disso, independentemente de um código em OpenCL ser suportado por uma grande variedade de dispositvos, isso não significa que o código será executado de forma otimizada em todos eles sem qualquer esforço da parte do programador [$CUDA^{TM}$ 2009b].

5. Conclusões

A utilização da computação paralela vêm sendo cada vez mais necessária para o processamento de grandes volumes de dados, contidos em vários tipos de problemas de diversas áreas da ciência. Com isso, a busca pelo aumento do desempenho computacional se torna significante à medida em que o volume de dados aumenta. Não menos importante, a utilização de um bom modelo de programação também se torna necessário para viabilizar o acesso e a computação dos dados. Modelos de programação como CUDA e OpenCL permitem que aplicações possam ser executadas mais facilmente na GPU.

Diversos trabalhos já utilizaram CUDA ou OpenCL para solucionar vários tipos de problemas paralelizáveis. Outros trabalhos apresentaram uma comparação de desempenho computacional entre ambos os modelos. Por exemplo, Karimi *et al.* (2010) apresentaram uma comparação de desempenho computacional entre CUDA e OpenCL. Eles mostraram que, apesar de o OpenCL fornecer um código portátil para execução em diferentes arquiteturas de GPUs, sua generalidade pode implicar em um baixo desempenho [Karimi et al. 2010]. Por outro lado, Fang *et al.* (2011) realizaram uma comparação de desempenho abrangente entre CUDA e OpenCL. Os resultados deles mostraram que, para a maioria das aplicações, CUDA se mostra, no máximo, 30% melhor do que o OpenCL [Fang et al. 2011]. Considerando apenas esses resultados, percebe-se que não há uma conclusão clara sobre qual arquitetura realmente é mais eficaz. Apenas é possível observar que, enquanto CUDA pode apresentar um bom desempenho computacional para a computação de algumas tarefas, o OpenCL pode se mostrar tão eficiente quanto CUDA.

Com base na revisão bibliográfica realizada, não foi encontrado algum artigo com-

pleto e recente que realiza uma comparação teórica, destacando claramente qual modelo, de fato, pode ser considerado mais adequado. Portanto, o objetivo deste trabalho foi comparar CUDA com o OpenCL apenas em relação à aspectos de $hardware$, $software$, tendências tecnológicas e facilidades de utilização. Apesar de ser uma tecnologia proprietária, foi possível concluir que CUDA, além de ser o primogênito e considerado mais "maduro", pode ser uma escolha mais viável em comparação com OpenCL.

A escolha do OpenCL pode parecer mais óbvia por ser possível desenvolver programas que poderiam ser executados em qualquer GPU, ao invés de desenvolver uma versão (em CUDA) para execução apenas nas placas da $NVIDIA^{®}$. Entretanto, na prática essa escolha pode ser um pouco mais complicada, já que o OpenCL oferece funções e extensões que são específicas para cada família [$CUDA^{TM}$ 2009b]. Ainda, por ter um modelo de gerenciamento para a portabilidade em multiplataformas e multifornecedores, o OpenCL pode ser considerado mais complexo [Kirk and Hwu 2011].

Trabalhos futuros poderão realizar comparações mais complexas entre CUDA e OpenCL. Por exemplo, aspectos como execução de instruções à nível de $hardware$ poderão ser analisados. Adicionalmente, possíveis novas arquiteturas e novos modelos de programação poderão surgir e serem investigados para a realização de estudos comparativos.

Referências

Barak, A., Ben-Nun, T., Levy, E., and Shiloh, A. (2010). A package for opencl based heterogeneous computing on clusters with many gpu devices. In *Cluster Computing Workshops and Posters (CLUSTER WORKSHOPS), 2010 IEEE International Conference on*, pages 1–7. IEEE.

$CUDA^{TM}$, N. (2007). *Accelerating MATLAB with CUDA*, volume 1. NVIDIA Corporation.

$CUDA^{TM}$, N. (2009a). *NVIDIA CUDA C Programming Best Practices Guide*. NVIDIA Corporation, 2701 San Tomas Expressway Santa Clara, CA 95050.

$CUDA^{TM}$, N. (2009b). *OpenCL Programming Guide for the CUDA Architecture*. NVIDIA Corporation.

$CUDA^{TM}$, N. (2013). *NVIDIA CUDA C Programming Guide*. NVIDIA Corporation, 2701 San Tomas Expressway Santa Clara, CA 95050, 5.0 edition.

Fang, J., Varbanescu, A. L., and Sips, H. (2011). A comprehensive performance comparison of cuda and opencl. In *Parallel Processing (ICPP), 2011 International Conference on*, pages 216–225. IEEE.

Gaioso, R. D. R. A., Jradi, W., de Paula, L. C. M., Alencar, W., Martins, W. S., Nascimento, H., and Caceres, E. (2013). Paralelização do algoritmo floyd-warshall usando gpu. In *XIV Simposio em Sistemas Computacionais*, pages 19–25.

Karimi, K., Dickson, N. G., and Hamze, F. (2010). A performance comparison of cuda and opencl. *arXiv preprint arXiv:1005.2581*.

Khronos, G. (2013). The open standard for parallel programming of heterogeneous systems. http://www.khronos.org/opencl/.

Kirk, D. B. and Hwu, W. (2011). *Programando para Processadores Paralelos*. Elsevier, 1 edition.

Komatsu, K., Sato, K., Arai, Y., Koyama, K., Takizawa, H., and Kobayashi, H. (2010). Evaluating performance and portability of opencl programs. In *The fifth international workshop on automatic performance tuning*.

Little, J. and Moler, C. (2013). Matlab gpu computing support for nvidia cuda-enabled gpus. http://www.mathworks.com/discovery/matlab-gpu.html.

Liu, X., Cheng, L., and Zhou, Q. (2013). Research and comparison of cuda gpu programming in matlab and mathematica. In *Proceedings of 2013 Chinese Intelligent Automation Conference*, pages 251–257. Springer.

Paula, L. C. M. (2013). Paralelização e comparação de métodos iterativos na solução de sistemas lineares grandes e esparsos. *ForScience: Revista Científica do IFMG*, 1(1):01–12.

Paula, L. C. M., Soares, A. S., Lima, T. W., Delbem, A. C. B., Coelho, C. J., and Filho, A. R. G. (2014). Parallelization of a modified firefly algorithm using gpu for variable selection in a multivariate calibration problem. *International Journal of Natural Computing Research*, 4(1):31–42.

Paula, L. C. M., Soares, A. S., Soares, T. W., Martins, W. S., Filho, A. R. G., and Coelho, C. J. (2013a). Partial parallelization of the successive projections algorithm using compute unified device architecture. In *International Conference on Parallel and Distributed Processing Techniques and Applications*, pages 737–741.

Paula, L. C. M., Souza, L. B. S., Souza, L. B. S., and Martins, W. S. (2013b). Aplicação de processamento paralelo em método iterativo para solução de sistemas lineares. In *X Encontro Anual de Computação*, pages 129–136.

Reese, J. and Zaranek, S. (2011). Gpu programming in matlab. http://www.mathworks.com/company/newsletters/articles/gpu-programming-in-matlab.html.

Simek, V. and Asn, R. R. (2008). Gpu acceleration of 2d-dwt image compression in matlab with cuda. In *Computer Modeling and Simulation, 2008. EMS'08. Second UKSIM European Symposium on*, pages 274–277. IEEE.

Smith, S. M. (2011). The gpu computing revolution.

Stallings, W. (2002). *Arquitetura e Organização de Computadores*. Prentice Hall, São Paulo, SP, Brasil, 5 edition.

Stone, J. E., Gohara, D., and Shi, G. (2010). Opencl: A parallel programming standard for heterogeneous computing systems. *Computing in science & engineering*, 12(3):66.

Suganuma, T., Krishnamurthy, R. B., Ohara, M., and Nakatani, T. (2013). Scaling analytics applications with opencl for loosely coupled heterogeneous clusters. In *Proceedings of the ACM International Conference on Computing Frontiers*, page 35. ACM.

Tsuchiyama, R., Nakamura, T., Iizuka, T., Asahara, A., Son, J., and Miki, S. (2010). *The OpenCL Programming Book*. Fixstars.

Protótipo de um sistema de comando por voz baseado em Arduino

Glauber R. Lemos, Tercio A. S. Filho

[1]Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Caixa Postal 75704-020 – Catalão – Goiás – Brasil

{glauber.ratti.l,tercioas}@gmail.com

Abstract. *This paper presents the development of a prototype of a voice recognition system capable of perfor features of a home automation system. The concern in this study was an implementation of low cost prototype. For such, an application for mobile devices was created focused on the Android platform (Client), which through Wi-Fi communicates with the Arduino (Server) to control electronics in a residence.*

Resumo. *Este artigo foi desenvolvido com o objetivo de apresentar um protótipo de um sistema de reconhecimento de voz capaz de realizar características de um sistema de automação residencial. A preocupação neste trabalho foi o desenvolvimento de um protótipo simples e de baixo custo. Para tal, foi criado um aplicativo para dispositivos móveis focado na plataforma Android (Cliente), que por meio de Wi-Fi se comunica com o Arduino (Servidor) para controlar eletroeletrônico em uma residência.*

1. Introdução

Com o rápido crescimento da tecnologia há sempre o interesse ou necessidade de torná-la presente em aspectos de nossas vidas, como uso pessoal, educacional e comercial. Automação Residencial, ou Domótica utiliza a tecnologia para possibilitar o monitoramento, controle e interação com dispositivos remotamente, sendo aplicada em casas, indústrias, escritórios e outros locais que se beneficiariam com vantagens desta tecnologia, que seriam uso eficiente de energia, conforto, segurança, praticidade, entre outros fatores.

A palavra Domótica é a junção da palavra latina *Domus* (casa) e do termo Robótica e o significado está relacionado à instalação de tecnologia em residências, com o objetivo de melhorar a qualidade de vida, aumentar a segurança e viabilizar o uso racional dos recursos para seus habitantes [Angel 1993]. Existem outras denominações para a Domótica, entre elas estão "Edifício Inteligente", "Casa Inteligente", "Ambiente Inteligente", "Automação Residencial", entre outros.

Domótica tem despertado atenção em famílias que querem adquirir os benefícios desta tecnologia. Não só pela comodidade, mas também para aspectos de acessibilidade [Bolzani 2004], permitindo que as tarefas domésticas sejam executadas de forma simples e acessível a pessoas idosas, pessoas com problemas de mobilidade ou deficiência física, oferecendo maior conforto e melhorando o seu nível de vida. Entretanto, a automação residencial ainda não recebeu uma ampla aceitação e atenção. Isso devido principalmente ao elevado custo de implementação e a complexidade. A Domótica seria mais acessível

se não houvesse a necessidade de utilizar tecnologias avançadas e complexas, e sim a utilização de dispositivos mais simples e de baixo custo.

No contexto de Domótica vários trabalhos vêm sendo realizados, como por exemplo: Al-Ali e Al-Rousan [Al-Ali & AL-Rousan 2004] apresentam o projeto e o desenvolvimento de um sistema de automação baseado em Java que pode monitorar e controlar eletrodomésticos através da *World Wide Web*; em Piyare e Tazil [Piyare & Tazil 2011] é apresentado um sistema Domótico de baixo custo, flexível e seguro de telefone celular baseado em comunicação *Bluetooth*; ElShafee e Hamed [ElShafee & Hamed 2012] apresentam um protótipo de um sistema de automação residencial utilizando a tecnologia *Wi-Fi*;

Baseado nas características de Domótica, este artigo apresenta-se um protótipo de um sistema que controla remotamente dispositivos elétricos de uma casa por comando de voz, com o intuito de realizar comandos similares aos sistemas mais sofisticados, com dispositivos de baixo custo e de um modo interativo. Com o protótipo é possível ligar e desligar lâmpadas, televisão, ventilador, e qualquer outro dispositivo eletroeletrônico da residência através de comando de voz. Os comando são feitos através de um dispositivo Android (Cliente) que se encontra conectado, via *Wi-Fi*, ao Arduino (Servidor) que liga e desliga tais dispositivos eletroeletrônicos.

A organização deste trabalho está dividido da seguinte forma: Na Sessão 2 apresenta-se conceitos para melhor entendimento do protótipo em geral; Na Sessão 3 trata-se do desenvolvimento do protótipo, de como o protótipo foi criado, já que foram esclarecidos alguns conceitos na Sessão 2; Posteriormente, na Sessão 4, serão apresentadas análises de testes realizados; E na Sessão 5 apresenta-se a conclusão e possíveis trabalhos futuros.

2. Conceitos

Para o melhor entendimento do protótipo proposto e deste trabalho em geral, nesta Sessão serão apresentados alguns conceitos relacionados aos dispositivos utilizados.

2.1. Arduino, Shield Ethernet

Arduino é uma plataforma baseada em hardwares e softwares livres, desenvolvida para desenvolver sistemas de forma simples, os quais interagem com o ambiente, utilizando entradas a partir de uma variedade de sensores (sensores de temperatura, luz, som e etc.) ou interruptores. Alguns exemplos de saídas desses sistemas são o controle de LEDs, motores, displays, auto falantes e outros periféricos [Arduino d].

O Arduino é composto por duas partes: o hardware, um conjunto de componentes eletrônicos montados numa placa de circuito impresso. E o software, o programa que será desenvolvido pelo programador e armazenado na memória *Flash*. Existe também a interface gráfica ou Ambiente de Desenvolvimento Integrado (IDE - *Integrated Development Environment*) que é aonde o programador vai criar seus programas e depois carregar para o hardware do Arduino. São os programas que gerenciam o hardware no que será realizado. [Silveira 2011].

O ambiente Arduino foi projetado para ser fácil de manipular, para iniciantes que não possui experiência em software ou eletrônica [Margolis 2011].

Além disso, o Arduino é uma plataforma relativamente barata e bastante flexível [Neto, Júnior, Neiva and Farinhaki 2010].

Existem diversos modelos de Arduino, cada um apropriado as necessidades do programador. Para o desenvolvimento do protótipo foi utilizado o Arduino Mega 2560 (Figura 1), devido a elevada quantidade de entradas e saídas em relação aos outros modelos, por apresentar uma memória *Flash* de 256 KB e com uma incrível potência de processamento [Arduino b].

O Arduino também permite encaixar em suas portas outras placas chamadas *shields*, com a finalidade de ampliar suas funcionalidades [Arduino a]. Existem vários tipos de *shields*, como por exemplo: para manipulação de motores, conexão por *blutooth*, sistemas de rede sem fio, entre outras. Para este protótipo foi utilizado a *Shield Ethernet* (Figura 1), no qual permite que o Arduino se conecte à internet ou uma rede Ethernet, possibilitando assim que o Arduino receba sinais de qualquer dispositivo Android [Arduino c].

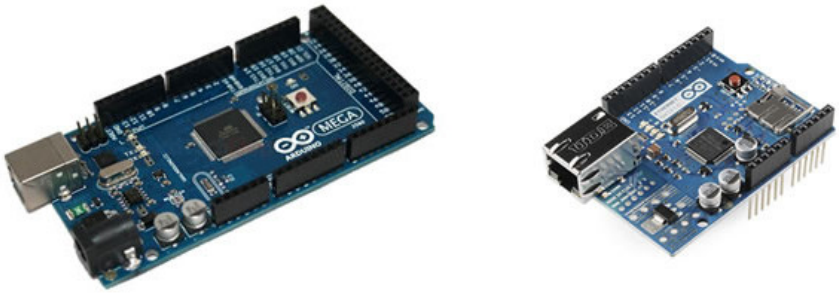


Figure 1. Dispositivos utilizados no protótipo. (a) Arduino Mega 2560; (b) *Shield Ethernet*

De forma sucinta, o Arduino é o elemento do protótipo capaz de receber dados da rede Ethernet e processa-los, e controlar ou monitorar transdutores conectados ao módulo Arduino. Os eletroeletrônicos são ligados ou desligados através do módulo relé, quando recebe algum sinal do Arduino.

2.2. Android e Android SDK

Android é um sistema operacional de código aberto baseado em Linux para dispositivos móveis desenvolvido pela Google e outras empresas, que juntas formam a *Open Handset Alliance* [Open Handset Alliance]. O Sistema operacional Android alcançou no terceiro trimestre de 2013 81,3% do mercado global de sistemas operacionais para smartphones. O segundo colocado apresenta uma porcentagem de 13,4%, tornando visível a popularidade da plataforma Android [Bicheno 2013]. Este fator de popularidade se deve ao fato de qualquer fabricante de hardware pode usar Android gratuitamente, desde que o Android apresente os serviços do Google como Gmail, Google Now, e Google Maps [Kovach 2013].

As aplicações para Android são criadas em uma linguagem de programação baseada no Java. Para auxiliar a criação de aplicações destinadas à Android, a Google disponibilizou a SDK (SDK - *Software Development Kit*) que possui um conjunto de ferramentas, tais como: uma API, um emulador de *smartphone* Android para realizar testes

das aplicações, códigos com exemplos, entre outras funcionalidades. A Google também disponibilizou um *plugin* para o IDE Eclipse, permitindo a integração entre este e o SDK, possibilitando que programadores crie aplicações do Android através do Eclipse.

A escolha de criar uma aplicação para Android no desenvolvimento deste protótipo, foi determinada pelo fato da plataforma Android ser utilizada na maior parte dos *smartphones* e pela facilidade de se programar em Java, já que esta é bem difundida. O IDE Eclipse foi utilizado para criar a aplicação por ser um IDE gratuito, pela possibilidade de programar aplicações para Android e por ser uma ferramenta já familiarizada pelos desenvolvedores do protótipo. Esta aplicação atuará como cliente, no qual receberá os comandos de voz do usuário os enviará ao servidor, isto é, o Arduino.

2.3. TCP e UDP

Segundo Tanenbaum [Tanenbaum 2003] a Internet tem dois protocolos principais na camada de transporte, um protocolo sem orientação a conexões e outro orientado a conexões. O protocolo sem conexões é o UDP (*User Datagram Protocol*) e o protocolo orientado a conexões é o TCP (*Transmission Control Protocol*).

O protocolo UDP, não realiza controle de fluxo, controle de erros ou retransmissão após a recepção de um segmento incorreto. O UDP é um protocolo simples e tem alguns usos específicos, como interações cliente/servidor e multimídia [Tanenbaum 2003]. Transmissões de áudio e vídeo são exemplos de aplicações UDP. Neste tipo de transmissão, a perda de uma pequena quantidade de dados não traz consequências graves ao processo. Porém, para a maioria das aplicações da Internet, é necessária uma entrega confiável e em sequência. O UDP não pode proporcionar isso, e assim foi preciso criar outro protocolo, o TCP.

O TCP foi projetado especificamente para oferecer um fluxo de bytes fim a fim de forma confiável [Tanenbaum 2003]. Este protocolo é responsável por enviar os dados da mesma forma que foram transmitidos e sem erros. Pelo fato de ser confiável, o protocolo TCP é usado em aplicações de rede (e-mail, telnet, ftp, etc), compartilhamento de arquivos, entre outros.

O protocolo utilizado neste protótipo foi o TCP, devido a confiança de que as mensagens enviadas do dispositivo Android chegará ao servidor, assim se houver algum erro com a mensagem enviada, esta será reenviada.

3. Desenvolvimento do Sistema

Como já citado, o propósito deste trabalho é apresentar um protótipo de um sistema com características da Domótica, que foi desenvolvido de forma simples e com dispositivos de baixo custo. Para facilitar o entendimento, a Figura 2 ilustra a estrutura do protótipo, no qual, ao enviar um comando de voz utilizando um dispositivo com a plataforma Android (cliente) e com a aplicação presente. O comando será enviado ao servidor (Arduino com *Shield Ethernet*) através da *Wi-Fi* pelo protocolo TCP. Ao receber o comando, o Arduino verifica o comando, que por sua vez realiza o controle ou monitoramento dos transdutores conectados ao Arduino.

No protótipo houve o desenvolvimento de dois *softwares*, um para criar o aplicativo para plataforma Android, que atuará como cliente, e outro para o Arduino que será

Figure 2. Esquema do protótipo desenvolvido

o servidor. Houve também o desenvolvimento relacionado ao *hardware* do Arduino e seus dispositivos complementares. Para um maior esclarecimento, a Sessão 3 foi divida da seguinte forma: Desenvolvimento relacionado ao Android, no qual apresenta como o aplicativo foi criado, e Desenvolvimento relacionada ao Arduino, que se trata tanto do desenvolvimento do *software* como a do *hardware* do Arduino e seus dispositivos complementares.

3.1. Desenvolvimento relacionada ao Android

No aplicativo para Android, foi desenvolvido um campo de inserção, onde o usuário deve digitar o número IP do servidor, e um botão que quando acionado, a aplicação estará pronta para receber o comando de voz do usuário, como apresentado na Figura 3. O sistema de reconhecimento de comando de voz é uma interface do Google, essa mesma interface pode ser encontrada em seu site de busca [Google].

3.2. Desenvolvimento relacionada ao Arduino

No desenvolvimento do *software* para o Arduino foi configurado o número IP do servidor, a máscara de rede e o *Gateway*. Ao receber uma mensagem do cliente, o *software* faz uma análise para verificar se essa mensagem está configurada para realizar algum comando. Caso a mensagem recebida não estiver configurada, o Arduino não fará nenhuma ação.

No desenvolvimento relacionada ao *hardware*, no Arduino foi acoplado a *Shield Ethernet* e conectado através de um cabo de rede com o roteador em uma das portas LAN (LAN - *Local Area Network*). No intuito de fornecer energia ao dispositivo e também para possibilitar que a programação desenvolvida seja transferida para o Arduino, este é ligado ao computador através de um cabo USB (USB - *Universal Serial Bus*).

Para realização dos testes foi utilizado um módulo relé, conectado entre uma das saídas do Arduino e no circuito elétrico da casa. Assim, quando o Arduino receber o comando, ele mandará o sinal correspondente para a determinada saída, controlando o módulo relé.

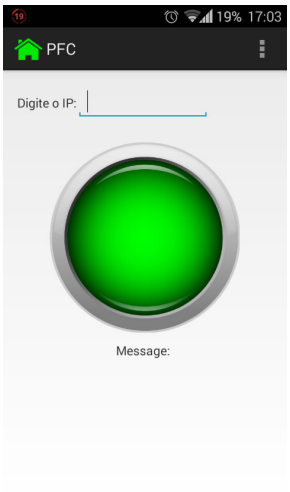


Figure 3. Aplicativo criado para plataforma Android

4. Análise e Testes

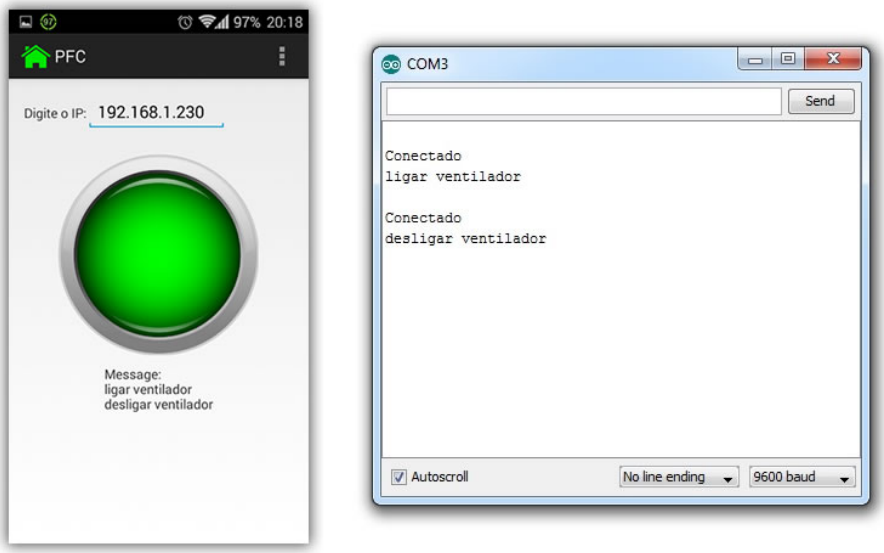
Neste protótipo foi utilizado um módulo relé, e este foi conectado a um ventilador. Outros dispositivos eletroeletrônico foram representados por LEDs de diferentes cores. A televisão foi representada por um LED de cor azul e a lâmpada por um LED de cor banca. Logo, o Arduino foi configurado para reconhecer 6 comando, que são: “ligar lâmpada”, “desligar lâmpada”, “ligar ventilador”, “desligar ventilador”, “ligar tv” e “desligar tv”.

Na Figura 4, estão presentes os resultados de um dos testes realizados, em que os comandos “ligar ventilador” e “desligar ventilador” foram reconhecidos pelo aplicativo quando o usuário emitiu o comando de voz e foram enviados ao Arduino. Os dados recebidos pelo Arduino podem ser visualizados com o auxílio da ferramenta *Serial Monitor* disponível no IDE do Arduino. E finalmente o ventilador é ligado e posteriormente desligado como consequência dos comandos emitidos.

Foram realizados 34 testes por 6 pessoas diferentes com conhecimento de como proceder com o aplicativo e com a instrução. Cada usuário foi instruído a emitir os 6 comandos existentes. Todos os testes foram satisfatórios, visto que todos os comando emitidos pelos usuário obtiveram as ações correspondentes. O aplicativo apresentou dificuldade em carregar a interface de reconhecimento de voz da Google nos dois primeiro teste. Tal problema é consequência do desempenho inferior do *smartphone* utilizado.

5. Conclusão e Trabalhos Futuros

Neste trabalho foi apresentado um protótipo de um sistema de reconhecimento de comando de voz, capaz de controlar eletroeletrônico de uma residência visando algumas características de um sistema Domótico.



(a) Comando emitido pelo usuário (b) Ferramenta *Serial Monitor*

Figure 4. Resultados de um dos testes realizados. (a) Reconhecimento dos comandos de voz emitidos pelo usuário; (b) Resultados do Arduino visualizados pela ferramenta *Serial Monitor*.

No protótipo foi utilizado um Arduino e outros dispositivos complementares (*Shield Ethernet* e módulo relé) que são tecnologias de baixo custo. Foi utilizado também o IDE Eclipse para programar, já que este é gratuito e possibilita a programação para a plataforma Android, que é o sistema operacional mais utilizado. Por esses motivos, foi possível criar o protótipo de forma simples e de baixo custo.

Os testes realizados tiveram resultados satisfatórios, visto que os comando emitidos pelos usuário obtiveram as ações correspondentes.

Para trabalhos futuros esse protótipo será otimizado, adicionando algumas funcionalidades a ele, como por exemplo o envio de mensagens do Arduino para o Android para confirmar as ações realizadas, e em caso de a mensagem recebida do Android não for reconhecida, informar que esta não existe. Outro exemplo seria a construção de um aplicativo para o *smartphone* iPhone.

References

Al-Ali, A. R. & AL-Rousan, M. (2004). Java-based home automation system. *IEEE Transactions on Consumer Electronics*, 50(2).

Open Handset Alliance. Oveview. Disponível em: <http://www.openhandsetalliance.com/android_overview>. Acesso em: 15 de Fevereiro de 2014.

Angel, P. M. (1993). *Introducción a la domótica*, volume 1. Escuela Basileño-Argentina de Informatica, EBAI.

Arduino. Arduino ethernet shield. Disponível em: <http://arduino.cc/en/Main/ArduinoEthernetShield>. Acesso em: 15 de Fevereiro de 2014.

Arduino. Arduino mega 2560. Disponível em: <http://arduino.cc/en/Main/arduinoBoardMega2560>. Acesso em: 15 de Fevereiro de 2014.

Arduino. Arduino shield. Disponível em: <http://arduino.cc/en/Main/ArduinoShields>. Acesso em: 15 de Fevereiro de 2014.

Arduino. What is arduino. Disponível em: <http://www.arduino.cc/en/Guide/Introduction>. Acesso em: 15 de Fevereiro de 2014.

Bicheno, S. (2013). Android captures record 81 percent share of global smartphone shipments in q3 2013. Disponível em: <http://blogs.strategyanalytics.com/WSS/post/2013/10/31/Android-Captures-Record-81-Percent-Share-of-Global-Smartphone-Shipments-in-Q3-2013.aspx>. Acesso em: 15 de Fevereiro de 2014.

Bolzani, C. A. M. (2004). *Residências Inteligentes*. Livraria da Física.

ElShafee, A. & Hamed, K. A. (2012). Design and implementation of a wifi based home automation system. *World Academy of Science, Engineering and Technology*, 6(8).

Neto, A. L. R., Júnior, A. M., Neiva, E. C. R. & Farinhaki, R. (2010). *Sistema de Medição de Campo Magnético Baseado no Efeito Hall e Arduino*. Monografia. Dep. Acadêmico de Eletrônica., Universidade Tecnológica Federal do Paraná, Curitiba.

Google. <http://www.google.com>.

Kovach, S. (2013). How android grew to be more popular than the iphone. Disponível em: <http://www.businessinsider.com/history-of-android-2013-8?op=1>. Acesso em: 15 de Fevereiro de 2014.

Margolis, M. (2011). *Arduino Cookbook*. O'Reilly Media.

Piyare, R. & Tazil, M. (2011). Bluetooth based home automation system using cell phone. *2011 IEEE 15th International Symposium on Consumer Electronics*.

Silveira, A. S. (2011). *Experimentos com o Arduino: Monte seus próprios projetos com o Arduino utilizando as linguagens C e Processing*. Ensino Profisional Editora, São Paulo, SP, 1 edition.

Tanenbaum, A. S. (2003). *Redes de computadores*. Campus (Elsevier), 4th edition.

Uma Abordagem sobre Segurança em Sistemas RFID

Rogéria Oliani[1], Alexandre César R. da Silva[1], Tércio Alberto dos Santos Filho[2]

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

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

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

Abstract. *Radio Frequency Identification (RFID) is a generic term for technologies that use radio waves to automatically identify people or objects, through the use of tags and readers. The diversity of applications that can be given to this technology are extensive. Thus, demand careful consideration in relation to security issues, which can range from loss of privacy, the large financial losses. This paper presents a study to RFID, with its main features and applications, as well as address issues of security, describing various types of attacks and countermeasures.*

Resumo. *Identificação por Rádio Frequência (RFID - Radio Frequency Identification) é um termo utilizado para tecnologias que utilizam a Rádio Frequência para identificação de objetos ou pessoas, através do uso de tags (etiquetas) e leitores. A diversidade de aplicações que podem ser dadas a essa tecnologia é extensa. Com isso, demanda uma análise cuidadosa em relação a questões de segurança, que podem abranger desde a perda da privacidade, a grandes prejuízos financeiros. Neste artigo apresenta-se um estudo ao RFID, com suas principais características e aplicações; bem como aborda questões de segurança, descrevendo diversos tipos de ataques e contramedidas.*

1. Introdução

Identificação por radiofrequência (RFID - *Radio Frequency Identification*) é um termo genérico para tecnologias que utilizam ondas de rádio para identificar automaticamente pessoas ou objetos. Um sistema RFID possui dois componentes básicos: Tag RFID, ou Etiqueta RFID, e Leitor RFID, ou Interrogador. A tag RFID, basicamente, possui um microchip ligado a uma antena, os quais são acondicionados em uma embalagem (encapsulamento) apropriada ao objeto ou pessoa a que se destina identificar (cartão de crédito, chave de veículo, prego para identificação de árvores ou paletes, etiquetas de vestuários, dentre outros). O Leitor RFID é um dispositivo utilizado para se comunicar com as tags através da emissão de ondas de rádio. Em relação a sua fonte de energia, as tags podem ser divididas em 3 tipos [RFID Journal Brasil 2013]:

1. Passiva: Não possui fonte de energia. A energia necessária ao seu funcionamento é recebida do leitor através dos sinais emitidos por este. Em virtude desta característica, são mais baratas e têm uma maior duração em comparação as tags ativas. Contudo, possuem capacidade computacional e memória limitada.

2. Semi-passiva ou semi-ativa: possuem bateria interna, porém utiliza a energia fornecida pelos leitores para transmitir o sinal a estes, como ocorre com as tags passivas. Neste caso, a bateria fornece energia ao seu microchip, permitindo que este tenha uma maior capacidade de processamento.
3. Ativas: possuem fonte de energia interna, o que possibilita o envio de sinais de transmissão de dados ao leitor, bem como alimentar circuitos mais complexos e sensores. Este tipo de tag possui um valor comercial mais alto, e um tempo de vida menor em comparação as tags passivas.

Outra forma de classificação das tags é dada pela frequência do sinal de transmissão de dados. Estas podem atuar em [Finkenzeller 2010]:

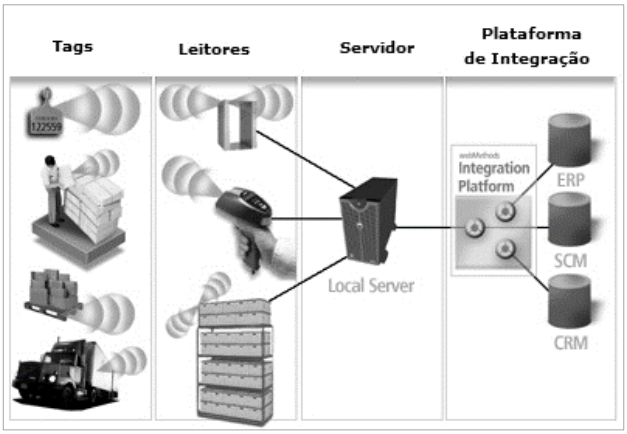
1. Baixa Frequência (LF – *Low Frequency*): atuam em uma frequência de 30 a 300 kHz, e possuem uma transferência de dados lenta, bem como um pequeno alcance de leitura (até 1 metro);
2. Alta Frequência (HF – *High Frequency*): a frequência encontra-se em uma faixa de 3 a 30 MHz. Elas geralmente podem ser lidas até 1 metro de distância, e possuem transmissão de dados mais rápida que as tags de baixa frequência, contudo, consomem mais energia do que estas.
3. Ultra Alta Frequência (UHF – *Ultra High Frequency*): atuam em uma faixa de frequência de 300 MHz a 3 GHz, e tipicamente operam entre 866 e 960 MHz. Tags UHF possuem taxas de transferência mais altas e maior alcance do que as tags de alta e baixa frequência. No entanto, as ondas de rádio, nesta frequência, não passam por itens com alto teor de água. Em comparação às tags de baixa frequência, as tags UHF são mais caras e utilizam mais energia.
4. Microwave: atuam em frequência acima de 3 GHz. Tags Microwave têm taxas de transferência muito altas e podem ser lidas a longas distâncias, contudo, elas usam uma grande quantidade de energia e são mais caras em comparação as demais.

Para armazenar e fazer o controle das informações que são lidas das tags pelos leitores, temos os servidores. Os servidores são computadores que armazenam o banco de dados das tags, e efetuam a comunicação com os leitores RFID através de uma rede (padrão IEEE 802.11, IEEE 802.15.4, dentre outros) ou, simplesmente, através de uma porta USB. As tags podem conter diversas informações, ou apenas um número de identificação (ID). Através do ID o servidor pode identificar a tag e obter as informações relacionadas a esta. Estas informações podem ser utilizadas pelas organizações para efetuar o controle de folha de pagamento, produção, inventário, controle de processos em linha de produção, análise de perfil, tendências, dentre inúmeras outras.

Na Figura 1, apresenta-se o esquema de um sistema RFID. Neste, a leitura dos dados de diversos tipos de tags – que identificam animais, caixas, paletes e caminhões – são realizadas por diferentes leitores; sendo o modelo de cada leitor adequado ao tipo de tag que deseja efetuar a leitura. Os dados são enviados a um servidor local, no qual encontram-se armazenadas as informações referentes a cada item identificado com a tag. As informações localizadas no servidor local são utilizadas em uma plataforma integrada com o Sistema Integrado de Gestão Empresarial (ERP - *Enterprise Resource Planning*), Gestão de Relacionamento com o Cliente (CRM - *Customer Relationship Management*) e Gestão da Cadeia Logística (SCM - *Supply Chain Management*). Desta forma, as

informações das tags lidas, podem ser obtidas em tempo real por toda a cadeia, podendo chegar até o cliente.

Figura 1 – Esquema de um sistema RFID.



Fonte: [RFIDBr 2008], adaptada pelo autor.

Há uma grande diversidade de sistemas que são implementados utilizando RFID, os quais se estendem pela cadeia de suprimentos, pedágios, transporte público, dentro outros. Os prejuízos financeiros e à privacidade, que a falta de segurança neste sistemas podem trazer, motiva o estudo e implementação de diversas técnicas e equipamentos que auxiliem na proteção destes.

Neste artigo apresenta-se alguns padrões RFID utilizados, exemplos de aplicações, bem como aborda questões relacionadas à segurança, descrevendo diversos tipos de ataques e contramedidas. Este artigo está organizado da seguinte forma: Na Sessão 2, serão apresentados alguns padrões utilizados na tecnologia RFID. Na Sessão 3, serão apresentadas diferentes tipos de aplicação da tecnologia RFID. Na Sessão 4, é aborda a questão da segurança e da privacidade em sistemas RFID, e apresenta diversos tipos de ataques, bem como contramedidas que podem ser utilizadas. Na Sessão 5, serão apresentadas as conclusões.

2. Padrões

Para que seja realizada a comunicação entre tags, leitores e servidor RFID é necessário a utilização de protocolos, os quais definem as regras de como esta será realizada. Estas regras definem, dentro outras questões, quais sinais são reconhecidos, como a comunicação é realizada, qual o significado dos dados recebidos das tags, quais dispositivos podem transmitir a cada tempo (resolvendo problemas de colisão). Assim, é importante a padronização destas regras, para que sistemas e equipamentos diversos sejam compatíveis, diminuindo o custo destes e facilitando sua implantação e disseminação. Neste âmbito, duas organizações se destacam: ISO (*International Standards Organization*) e a EPC Global. A ISO é uma união mundial de instituições nacionais de normalização, tais como DIN (Alemanha) e ANSI (EUA) e contribui com inúmeros comitês e grupos de trabalho para o desenvolvimento de padrões de RFID [Finkenzeller 2010]. Dentre os padrões ISO podemos citar os de baixa frequência, utilizado no rastreamento de animais (ISO 11784, ISO 11785, ISO 14223), os de alta frequência utilizados em cartões inteligentes (ISO 10536, ISO 14443, ISO 15693) e os da série ISO 18000, utilizados no gerenciamento de itens, os quais atuam em diferentes

frequências. A EPC Global é uma organização sem fins lucrativos que visa a padronização do RFID através do Código Eletrônico de Produto (EPC - *Electronic Product Code*) e da EPC *Network*. O EPC é um meio para identificar de forma única paletes, caixas ou itens. Uma tag EPC não carrega informações pessoais. Todas as informações sobre o objeto com a tag EPC é administrada exclusivamente no EPCglobal Network. O EPCglobal *Network* é uma tecnologia, a qual permite a parceiros comerciais documentar e determinar a localização de bens individuais na cadeia de abastecimento, se possível em tempo real [Finkenzeller 2010]. Dentre os padrões EPC podemos citar o *Class* 0, *Class* 1 e o *Class* 1 *Gen* 2; sendo este último reconhecido pela ISO como padrão internacional (ISO 18000-6C).

3. Aplicações

Existe um grande número de sistemas que são implementados utilizando RFID. A diversidade de aplicações se estende pela cadeia de suprimento, agricultura, identificação de animais, controle de acesso, transporte público, pedágios, dentre outros.

A identificação por rádio frequência nas cadeias de suprimentos permite rastrear todo o processo produtivo, materiais utilizados na fabricação, inspeção, faturamento, distribuição, até chegar ao revendedor ou mesmo consumidor final. Na fábrica de automóveis da Volkswagen, na Eslováquia, seus veículos montados nas estações de serviços finais e processos de inspeção na unidade de Bratislava, são rastreados utilizando um sistema de localização em tempo real. A solução que emprega tecnologia RFID com o uso de tags ativas, permite que a empresa localize os veículos estacionados e identifique quando um carro entra ou sai de cada um dos vários processos, o que torna possível melhorar a eficiência dos estágios de produção final [Swedberg 2012].

No Brasil, o Departamento Nacional de Trânsito (Denatran) está em processo de implantação do SINIAV (Sistema Nacional de Identificação Automática de Veículos), o qual tem por objetivo aperfeiçoar a gestão do tráfego e a fiscalização de veículos, através do rastreamento destes utilizando tecnologia RFID [Perin 2013]. O projeto do Denatran determina a adoção obrigatória da tecnologia de identificação por radiofrequência em toda a frota brasileira de veículos, estimada atualmente em 85 milhões de unidades. Os primeiros testes começaram em outubro de 2012, utilizando placas convencionais associadas as tags RFID. Em seguida foram testados os sistemas de monitoramento de veículos emplacados com tags semi-ativas, chamadas PIVEs (Placas de Identificação veicular). Em virtude da evolução do padrão EPC Gen2 v2, considera-se no futuro o uso de tags passivas, que reduziriam os custos das PIVEs, porque não exigem emprego de baterias e são mais simples [Perin 2014].

A diversidade de aplicações do RFID, e a importância destas para as organizações que as implantam, tornam necessário que se dê atenção a segurança destes sistemas.

4. Segurança e privacidade

A segurança dos sistemas que utilizam RFID engloba questões relacionadas a proteção dos dados contidos nas tags e disponibilidade dos serviços ao qual se destina a identificação fornecida por estas. Além do número de identificação, as tags podem armazenar outros dados relacionados a “quem” ou a “o que” ela se destina identificar. Estes dados podem comprometer a segurança dos processos envolvidos, bem como comprometer a privacidade das pessoas que a utilizam.

A questão da privacidade em sistemas RFID incluem o vazamento de informações contidas nas tags, bem como o rastreamentos destas. As tags podem responder aos interrogadores sem o conhecimento de quem as esta portando ou de seus proprietários. Quando o número de identificação da tag é relacionada a dados pessoais, o problema se torna maior; pois permite, por exemplo, que um comerciante trace o perfil do consumidor utilizando rede de leitores tanto dentro, quanto fora do estabelecimento comercial. No Brasil, a privacidade tem sido discutida com relação a implantação do Sistema Nacional de Identificação Automática de Veículos (SINIAV). A Ordem dos Advogados do Brasil questiona o fato do sistema permitir conhecer a exata localização do veículo de uma pessoa, ferindo, assim, o direito constitucional à garantia de privacidade dos cidadãos [Leitão 2012].

Como visto, a privacidade para ser ferida não precisa necessariamente que o sistema de RFID sofra um ataque, pois a própria entidade que disponibilizou a tag para o usuário pode utilizar-se do conhecimento das informações contidas nesta, para interesses próprio.

4.1. Ataques

Os sistemas RFID são suscetíveis a ataques como todos os sistemas que envolvem transmissão e armazenamento de dados. Os objetivos de cada ataque podem ser muito diferentes; sendo, assim, é importante identificar os potenciais alvos para compreender os possíveis ataques. A seguir são apresentados alguns tipos de ataque que sistemas RFID podem sofrer:

1. *Eavesdropping* (espionagem): é um ataque passivo no qual um atacante escuta a comunicação entre a tag e o leitor [Boli, Simplot-Ryl e Stojmenovic 2010]. Um exemplo deste tipo de ataque é a obtenção de dados de um cartão de crédito, como: nome do proprietário, número do cartão, data de expiração, tipo de cartão; através da captura das transmissões realizadas entre um leitor de cartão de crédito e um cartão de crédito RFID. Os dados obtidos neste tipo de ataque podem ser utilizados em ataques mais complexos, como: *Replay* e *Tracking*.
2. *Man in the middle*: O atacante se posiciona em um local intermediário entre o leitor e a tag que se encontra fora do alcance de leitura do mesmo. Assim, interrompe o caminho de comunicação, e manipula as informações que serão transmitidas tanto para o leitor, como para a tag, engando os dois componentes [Bienert e Schalk 2013].
3. *Tracking*: O rastreamento utiliza dados contidos nas tags para identificar a presença destas em um determinado ambiente físico, sem a autorização de quem a porta ou de seu proprietário [Boli, Simplot-Ryl e Stojmenovic 2010]. Desta forma é possível identificar a trajetória do objeto ou da pessoa a ela associada. Mesmo que em uma tag esteja armazenado apenas seu número de identificação, seria possível em uma compra, por exemplo, a loja estabelecer em seu banco de dados um vínculo entre o número de identificação da tag e o cliente que adquiriu o produto no qual a tag se encontra. Desta forma, seria possível identificar a presença do cliente na loja, quando ele voltasse portando o objeto anteriormente adquirido.
4. *Replay*: neste tipo de ataque os dados transmitidos entre a tag e o leitor são capturados e, posteriormente, reutilizados de forma a forjar uma nova comunicação [Bienert e Schalk 2013].

5. *Cloning*: Os dados de uma tag válida são capturados pelo leitor do atacante e, posteriormente, escritos em uma outra tag (clone) [Finkenzeller 2010]. Desta forma a tag clonada se comportará como a original perante o leitor.
6. *Spoofing*: o invasor simula uma identidade diferente da que ele tem, no caso, ele simula uma tag válida, e assim, pode fazer uso de todos os privilégios que aquela tag proporciona. A diferença entre este tipo de ataque e o *Cloning* é que neste último há uma reprodução física de uma tag original, enquanto no ataque *Spoofing* é utilizado um equipamento eletrônico para emular ou imitar a tag original [Tehranipoor e Wang 2012].
7. *Denial of Service* (DoS): Os ataques de negação de serviço têm o objetivo de impedir que usuários legítimos consigam utilizar o sistema. Ataques DoS podem ser realizados, por exemplo, utilizando dispositivos que emitam sinais de ruído na faixa de frequência utilizada pela rede RFID, reduzindo a taxa de transferência e, consequentemente, emperrando o sistema [Xiao, Gibbons e Lebrun 2009]. Um outro exemplo seria a utilização não autorizada do comando *KILL*, desta forma as tags deixariam de responder aos leitores [Tehranipoor e Wang 2012].

Alguns ataques, em um primeiro momento, podem parecer não trazer prejuízos, como é o caso do *Eavesdropping*, contudo, servem como base para outros mais complexos (*Replay*, *Tracking*). As consequências geradas pelos diversos tipos de ataques podem atingir diferentes níveis de prejuízo financeiro, ou até mesmo relativos a privacidade de indivíduos, como podem ocorrer com o *Tracking*. Ataques como *Cloning* e *Spoofing* podem permitir acesso a áreas não autorizadas de empresas e residências, no caso destas utilizarem fechaduras com tecnologia RFID, bem como qualquer outro tipo de privilégio que se poderia ter com uma tag original.

A seguir, apresenta-se algumas contramedidas que podem ser utilizadas de forma a inibir diferentes tipos de ataques.

4.2. Contramedidas

Ataques ao sistema RFID podem causar grandes prejuízos financeiros as empresas que o utilizam, bem como a usuários finais, neste último caso, podendo atingir também a privacidade destes. Assim, faz-se necessário que contramedidas sejam tomadas para evitar ataques RFID. Dentre as várias contramedidas, podemos citar:

1. RSA *Blocker Tags*: é um produto desenvolvido pelos cientistas do laboratório RSA em conjunto com o Professor Ronald Rivest para a proteção da privacidade de consumidores. O RSA *Blocker Tag* cria uma região física a sua volta, a qual impede que os leitores de RFID singularizem as tags que se encontram nesta região [Juels, Rivest e Szydlo 2003].
2. *Kill Command*: é uma forma de proteger a privacidade dos consumidores enviando um comando para matar a tag, não sendo mais possível que esta seja lida por qualquer leitor RFID [Xiao, Gibbons e Lebrun 2009], [López 2008]. Desta forma, pode ser dada ao consumidor, por exemplo, a oportunidade de matar a tag antes de sair de uma loja, evitando o seu rastreamento.
3. Gaiola de Faraday: baseado na Gaiola de Faraday, é uma forma de bloquear as frequências de rádio utilizando um isolamento, o qual impede que os sinais do interior do objeto que possui esse isolamento alcance o seu exterior e vice-versa. Este isolamento pode ser simplesmente feito com folhas de metal, ou até mesmo, ser adquirido no mercado objetos como carteiras, bolsas, porta cartão, dentre

outros, confeccionados com material próprio para esta função. Contudo, este mesmo tipo de isolamento, poderia ser utilizado, por exemplo, para efetuar furtos de produtos em lojas, inibindo a leitura das tags dos produtos.

4. Criptografia: pesquisadores têm propostos várias versões de criptografias para serem utilizadas em sistemas RFID [López 2008], [Sun e Zhong 2012], [Noman, Rahman e Adams 2011], [Sharaf 2012], [Dong, Zhan e Wei 2013]. Criptografias podem ajudar a proteger o sistema contra diversos tipos de ataque, como: *Man in the middle* [Xiao, Gibbons e Lebrun 2009], *Cloning* [López 2008], *DoS* [Dong, Zhan e Wei 2013], dentre outros. O grande desafio têm sido a criação de criptografias leves o bastante para serem utilizadas em tags de baixo custo, eis que estas possuem capacidade computacional limitada (armazenamento, circuitos e consumo de energia) [López 2008].
5. *RFID Guardian*: é um dispositivo portátil alimentado por bateria que atua como um mediador das intereações entre os leitores e tags RFID. O Guardião RFID contém recursos de um leitor de RFID e de emulação de tag, os quais lhe permitem auditar e controlar as atividades de RFID [Cengage Learning 2010].

As contramedidas voltadas ao sistema RFID são aplicadas de acordo com o objeto ou pessoa que se deseja identificar. O nível de segurança e quanto se deseja investir financeiramente nesta, vai depender do valor do objeto a que se destina, e das informações que são armazenadas na tag ou aquelas a que esta permite o acesso no sistema. Assim, há contramedidas que podem ser utilizadas tanto por organizações, quanto por usuários finais, como é o caso das carteiras baseada na Gaiola de Faraday.

Dentre as contramedidas destacam-se as criptografias, em virtude destas ajudarem a proteger o sistema de diversos tipos de ataques. Justificando, assim, o grande empenho tido por diversos pesquisadores no estudo destas.

5. Conclusão

A diversidade de aplicações que são implementadas utilizando sistemas RFID vêm aumentando em todo o mundo. Ampliando, assim, o interesse de pesquisadores na busca de soluções para problemas de segurança, com o estudos de novas criptografias, principalmente, voltadas a aplicação em tags de baixo custo, as quais são as mais utilizadas. Há no mercado a oferta de diversos tipos de tags, com diferentes preços, criptografias, frequências, e outras características; devendo serem estas escolhidas de acordo com a sua aplicação, levando em consideração o custo x benefício. Tags ativas possibilitam a implementação de uma maior segurança, em virtude de sua capacidade computacional ser melhor em relação as passivas. Contudo, seu custo é mais alto, podendo representar uma parcela significante do custo de produção de produtos de baixo valor comercial.

Bibliografia

Bienert, R.; Schalk, G. H. RFID - MIFARE and Contactless Smartcards in Application. United Kingdom: BV, Elektor International Media, 2013.

Boli, M.; Simplot-Ryl, D.; Stojmenovic, I. RFID SYSTEMS - RESEARCH TRENDS AND CHALLENGES. United Kingdom: John Wiley & Sons, 2010.

Cengage Learning. RFID Hacking. In: _____ Ethical Hacking and Countermeasures: Linux, Macintosh and Mobile Systems. New York: [s.n.], 2010.

Dong, Q.; Zhan, ; Wei, L. A SHA-3 Based RFID Mutual Authentication Protocol and Its Implementation. 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC), 5-8 ago. 2013. 1 - 5.

Finkenzeller, K. RFID Handbook. United Kingdom: Wiley, v. 3, 2010.

Juels, A.; Rivest, L.; Szydlo, M. The Blocker Tag: Selective Blocking of RFID Tags for Consumer Privacy. 10th ACM conference on Computer and communications security. New York: [s.n.]. 2003.

Leitão, T. Sistema de identificação de veículos divide opiniões de especialistas, 03 outubro 2012. Disponivel em: <http://agenciabrasil.ebc.com.br/noticia/2012-10-03/sistema-de-identificacao-de-veiculos-divide-opinioes-de-especialistas>.

López, P. P. Lightweight cryptography in radio frequency identification (RFID) systems. Ph.D. THESIS, Universidad Carlos III de Madrid, Leganés,Espanha, 04 nov. 2008. Disponivel em: <http://e-archivo.uc3m.es/handle/10016/5093>.

Noman, A. N. M.; Rahman, S. M. ; Adams,. Improving Security and Usability of Low Cost RFID Tags. Ninth Annual International Conference on Privacy, 2011.

Perin, E. Ceitec inicia produção em volume do chip para o Siniav. RFID Journal Brasil, 18 novembro 2013. Disponivel em: <http://brasil.rfidjournal.com/noticias/vision ?11193/1>. Acesso em: 25 nov. 2013.

Perin, E. Siniav pode ser implantado em breve. RFID Journal Brasil, 06 fevereiro 2014. Disponivel em: <http://brasil.rfidjournal.com/noticias/vision?11415/1>. Acesso em: 27 fev. 2014.

RFID Journal Brasil. RFID Journal Brasil, 2013. Disponivel em: <http://brasil.rfidjournal.com>. Acesso em: 19 nov. 2013.

RFIDBr. Funcionamento RFID. RFIDBr, 28 novembro 2008. Disponivel em: <http: //www.rfidbr.com.br/index.php/funcionamento-rfid.html>. Acesso em: 01 mar. 2014.

Sharaf, M. RFID Mutual Authentication and Secret Update Protocol for Low-cost Tags. 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 25-27 jun. 2012. 71 - 77.

Sun, D.-Z.; Zhong, J.-D. A Hash-Based RFID Security Protocolfor Strong Privacy Protection. IEEE Transactions on Consumer Electronics, v. 58, n. 4, p. 1246-1252, 2012.

Swedberg, C. Volkswagen ganha eficiência no processo de acabamento de carros. RFID Journal Brasil, 18 junho 2012. Disponivel em: <http://brasil.rfidjournal.com/estudos-de-caso/vision?9626>. Acesso em: 25 fev. 2014.

Tehranipoor, M.; Wang, C. Security for RFID Tags. In: _____ Introduction to Hardware Security and Trust. New York: Springer, 2012. p. 283-302.

Xiao, Q.; Gibbons, T.; Lebrun, H. RFID Technology, Security Vulnerabilities, and Countermeasures. Supply Chain the Way to Flat Organisation, Publisher-Intech, p. 357-382, jan. 2009.

Sistema para Realização de Exercícios Fisioterapêuticos utilizando Realidade Virtual e Aumentada por meio de Kinect e Dispositivos Móveis

Flávia Gonçalves Fernandes[1], Sara Cristina Santos[1], Luciene Chagas de Oliveira[1], Mylene Lemos Rodrigues[1], Stéfano Schwenck Borges Vale Vita[1]

[1]Instituto de Engenharia e Tecnologia – Universidade de Uberaba (UNIUBE)
CEP: 38.408-343 – Uberlândia – MG – Brasil

flavia.fernandes92@gmail.com, saracristina33@hotmail.com,
luciene.oliveira@uniube.br, mylene.rodrigues@uniube.br,
stefanoborges@gmail.com

Abstract. *It's been some time that digital games are no longer seen as a form of entertainment harmful. Games are fast becoming an important tool to improve the treatment of patients ranging from those who are going through a serious illness like cancer or AIDS, even the demanding lighter procedures such as physiotherapy. This paper presents an application using Virtual and Augmented Reality (AVR) in the development of fisiogames, physiotherapy treatments applied in games through Microsoft Kinect process. In addition, the system also provides a mobile version, which provides a dual interactivity.*

Resumo. *Já faz algum tempo que jogos digitais deixaram de ser vistos como uma forma de entretenimento prejudicial à saúde. Games estão rapidamente se tornando uma ferramenta importante para melhorar o tratamento dos pacientes que vão desde aqueles que estão atravessando uma grave enfermidade, como o câncer ou a aids, até os que demandam procedimentos mais leves, como a fisioterapia. Este trabalho apresenta uma aplicação utilizando Realidade Virtual e Aumentada (RVA) no processo de desenvolvimento de fisiogames, jogos aplicados em tratamentos fisioterapêuticos por meio do Microsoft Kinect. Além disso, o sistema também apresenta uma versão mobile, o que proporciona uma dupla interatividade.*

1. Introdução

No mundo globalizado e tecnológico em que se vive, há sempre a necessidade de buscar mais inovações com a finalidade de facilitar e melhorar a vida da sociedade. A Realidade Virtual (RV) e Realidade Aumentada (RA) são tecnologias de interface do usuário extremamente promissoras e cada vez mais viáveis. Os avanços tecnológicos têm causado modificações significativas no ensino com a utilização cada vez maior de computadores, de softwares educativos e Internet, estes constituem pontos centrais em todo debate sobre o emprego de novas tecnologias em diversas áreas, como saúde, educação, comércio e entretenimento [Cardoso, 2007].

A maioria dos jogos eletrônicos exigem baixo tempo de resposta, reflexos aguçados, movimentos bruscos com alta velocidade de execução que são contraindicados em fases iniciais do processo de reabilitação. Esta característica pode ser prejudicial no

processo de cicatrização da lesão, aumentando inclusive o risco de produzir novas lesões [Costa, 2009].

Os fisiogames, em geral, são jogos produzidos focando-se no ganho de amplitude de movimentos, ganho de coordenação, força, resistência e precisão do movimento de modo a acelerar o processo de reabilitação. A realidade virtual recria a sensação de realidade para um indivíduo, possibilitando assim um tratamento motivador e mais adaptado ao paciente [Lamounier, 2004].

Partindo dessas premissas, foi utilizado o Microsoft Kinect como plataforma para desenvolvimento de um jogo que possa ser usado como ferramenta de apoio no processo de reabilitação fisioterapêuticas. O Microsoft Kinect é um sensor baseado na captura de movimento que utiliza o conceito de NUI (Interface Naturais). Este tipo de interface está focada na utilização de uma linguagem natural para interação humana com o aplicativo, como gestos, poses e comando de voz com um grau de precisão avançado [Gomes, 2006].

Nesta perspectiva, o objetivo deste trabalho é mostrar a implementação de uma aplicação utilizando RVA voltada para o tratamento fisioterapêutico, a qual consiste num jogo composto pela realização de uma sequência de exercícios terapêuticos pré-definidos no início do jogo, onde o paciente é estimulado a imitar os movimentos do personagem animado virtualmente. O Microsoft Kinect é utilizado como hardware para captura de movimentos do paciente no jogo. Em tempo de execução do jogo, os exercícios executados pelo paciente são comparados com a velocidade de execução e movimentos do exercício realizado pelo personagem virtual, então atribui-se uma pontuação ao paciente conforme o número de acertos do exercício. Além disso, foi desenvolvida uma versão deste sistema para dispositivos móveis com a finalidade de propiciar maior facilidade e interesse aos usuários.

2. Fundamentos Teóricos

Realidade Virtual (RV) é o uso de diversas tecnologias digitais para criar a ilusão de uma realidade que não existe de verdade, fazendo a pessoa mergulhar em mundos criados por um computador, ou seja é uma interface avançada para aplicações computacionais, onde o usuário pode navegar e interagir, em tempo real, em um ambiente tridimensional gerado por computador, usando dispositivos multissensoriais [Kirner, 2005].

Realidade Aumentada (RA) é a sobreposição de objetos virtuais gerados por computador em um ambiente real, utilizando para isso algum dispositivo tecnológico [14]. A Realidade Aumentada proporciona ao usuário uma interação segura e agradável, eliminando em grande parte a necessidade de treinamento, pelo fato de trazer para o ambiente real os elementos virtuais, enriquecendo e ampliando a visão que ele tem do mundo real. Para que isso se torne possível, é necessário combinar técnicas de visão computacional, computação gráfica e realidade virtual, o que origina como resultado a correta sobreposição de objetos virtuais no ambiente real [Azuma, 2001].

O Kinect é uma tecnologia de realidade virtual que é uma combinação de hardware e software contida dentro do sensor do Kinect. O sensor possui um hardware que oferece diversos recursos para auxiliar no processo de reconhecimento de gestos e voz, os principais são: Emissor de luz infravermelho, sensor RGB, sensor infravermelho, eixo motorizado e um conjunto de microfones dispostos ao longo do sensor. Existem várias tecnologias para capturar o movimento humano, tanto para animação quanto para aplicações de realidade virtual, neste trabalho utilizamos o Kinect para fazer o

monitoramento dos movimentos do paciente e para fazer a animação do personagem virtual [Harma, 2003].

Microsoft XNA é um framework gratuito e robusto com interface amigável desenvolvido pela Microsoft para criar jogos tanto para PC, console XBox 360 e Windows Phone 7. XNA foi projetado pensando nas pessoas que querem fazer seus próprios jogos e acham complicado trabalhar com DirectX, OpenGL e/ou outras APIs. Essa plataforma de programação gráfica, intitulada Microsoft XNA Game Studio, funciona como um meio de conexão entre as APIs do DirectX e o programador, possuindo uma série de funcionalidades e rotinas previamente compiladas, facilitando ao máximo o trabalho com efeitos e geometria espacial [Birck, 2007].

3. Metodologia

Durante o processo de desenvolvimento do fisiogame, foi feito um estudo detalhado sobre o funcionamento do sensor Kinect e como funciona o reconhecimento de esqueleto do sensor. Depois, foi realizada uma pesquisa das tecnologias necessárias para o desenvolvimento do jogo. Posteriormente, uma fisioterapeuta habilitada foi consultada sobre os tipos de movimentos que o paciente precisa realizar durante o processo de reabilitação motora da coluna lombar.

Em relação aos aspectos metodológicos e tecnológicos, para a implementação desta aplicação foi utilizada RVA por meio do desenvolvimento de ambientes virtuais, incluindo interações e animações, com uso da linguagem de programação C#. Além disso, foram utilizados o sensor Kinect e a plataforma XNA, ambos da Microsoft, visando criar para o usuário a possibilidade de interagir com ambientes virtuais atrativos, que facilitem a realização dos exercícios fisioterapêuticos de maneira correta.

Para o sensor Kinect, existem dois tipos de movimentos que são utilizados a partir do esqueleto do paciente: poses e gestos. Pose é uma forma de manter o corpo parado por um determinado tempo até que isso tenha algum significado. Gesto é o movimento do corpo em um espaço de tempo.

Para detectar o tipo de gesto, foi necessário implementar um algoritmo para rastrear os movimentos do usuário e comparar com os movimentos do exercício que foi criado via software. Para isso, primeiramente, foi criado uma classe para rastrear e identificar gestos. Esses gestos podem estar em três estados de rastreamento: não identificado, em execução e identificado.

Todo e qualquer tipo de movimento envolve tempo de execução e para que uma pose tenha validade é necessário que ela esteja sendo feita por um período de tempo. A etapa que avalia a pose ao longo do tempo chama-se rastreamento e a etapa que ocorre quando a pose é reconhecida como válida chama-se identificação.

Para fazer a implementação de um gesto neste fisiogame, foi preciso usar uma abordagem um pouco complexa, pois para fazer a detecção de um gesto usou-se uma abordagem que compara além da relação entre duas articulações, utiliza o ângulo entre três articulações, que define, por exemplo, o quão aberto deve estar o braço do paciente. Para calcular o ângulo entre três articulações utilizamos um método conhecido como produto escalar, apresentado na Figura 1. Este método é utilizado para calcular ângulos entre dois vetores (V e W) em um espaço de três dimensões.

$$\cos^{-1}\left(\frac{V.W}{\|v\|.\|w\|}\right)$$

Figura 1: Fórmula do produto escalar [Birck, 2007].

Nesta linha de raciocínio, foi realizada a detecção de cada gesto, também foi preciso levar em consideração uma margem de erro no valor de cada ângulo, pois comparações entre as articulações não devem ser feitas de forma exata, ou seja, é inviável exigir do paciente que os gestos sejam exatamente iguais.

A plataforma XNA foi escolhida para o desenvolvimento do fisiogame por ela oferecer um ambiente de desenvolvimento rico, de fácil aprendizagem e que funciona totalmente em ambiente de execução gerenciado. Ela é a interface que permite a interação entre o paciente (jogador) no jogo e o Kinect.

Além disso, XNA possui várias camadas que facilita a lógica de implementação do fisiogame. Na Figura 2 abaixo, pode ser visualizado o fluxograma básico do funcionamento da aplicação em XNA.

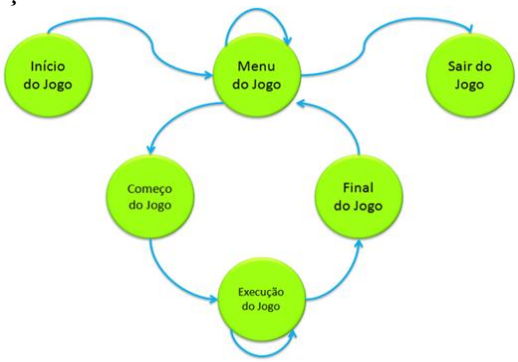


Figura 2: Fluxograma da lógica de funcionamento do fisiogame.

O XNA possui um gerenciador de dispositivo gráfico que permite lidar diretamente com a camada gráfica do dispositivo. Ele inclui métodos, propriedades e eventos que permitem consultar e alterar essa camada. Trocando em miúdos, é o gerenciador que lhe dará acesso aos recursos da placa de vídeo.

O gerenciador de conteúdo é um dos recursos mais interessantes do XNA. A forma com que foi desenvolvido permite que importemos qualquer tipo de conteúdo de ferramentas diferentes para o jogo. Em jogos que não utilizam XNA, os desenvolvedores precisam se preocupar em como carregar tais conteúdos, onde esse conteúdo está armazenado, se existem as bibliotecas corretas para importação e execução e mais uma série de dores de cabeça desnecessárias.

A lógica central do fisiogame inclui a preparação do ambiente no qual ele será executado, a execução do jogo em um loop até que o critério de fim de jogo seja alcançado e, finalmente, a limpeza desse ambiente. Estas funcionalidades são demonstradas a partir das seguintes classes em linguagem de programação C#:

- Game1() – Inicialização geral (Game1.cs);
- Initialize() – Inicialização do jogo (Game1.cs);

- LoadContent() – Inicializa e carrega recursos gráficos (Game1.cs);
- Run() – Inicia o loop do jogo (Program.cs). A cada loop:
- Update() – Captura os comandos do jogador, realiza cálculos e testa o critério de fim de jogo (Game1.cs);
- Draw () – Desenha os gráficos em tela (Game1.cs);
- UnloadContent() – Libera os recursos gráficos.

A maior parte do processamento do jogo acontece dentro do loop. É no loop que o jogo verifica os comandos do jogador, processa-os e dá o feedback no personagem, calcula a inteligência artificial, os movimentos são calculados e executados, as colisões entre objetos detectadas, a vibração do controle é ativada, o som é tocado, os gráficos desenhados na tela e os critérios para alcançar o fim do jogo checados. Praticamente tudo ocorre no loop do jogo. A cada loop as seguintes funções são inicializadas:

- Capturar os comandos do jogador;
- Executar os cálculos necessários (Inteligência Artificial, movimentos, detecção de colisões etc.);
- Verificar se o critério de fim de jogo foi alcançado – em caso positivo, o loop para;
- Desenhar gráficos em tela, gerar sons e respostas aos comandos do jogador;
- Finalizar os gráficos, dispositivos de entrada e som.

O jogo finaliza quando o paciente termina a execução de todos os exercícios selecionados anteriormente, e, a cada exercício executado de maneira e velocidade corretas, são atribuídos pontos ao paciente. Quanto maior a pontuação, melhor é a evolução do tratamento do paciente.

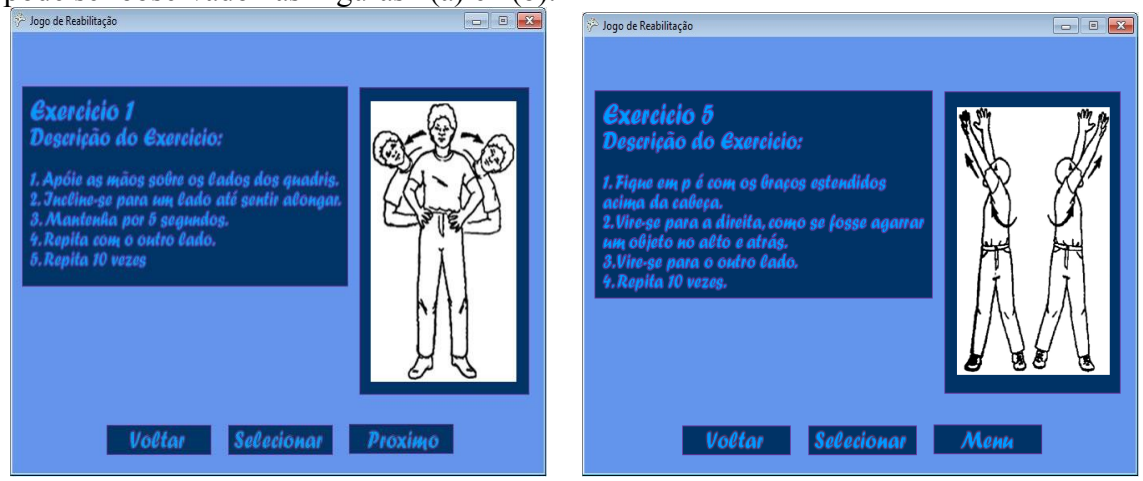
4. Resultados e Discussões

A partir das pesquisas conduzidas, foi realizada a implementação do fisiogame: um jogo utilizando Realidade Virtual e Aumentada que visa auxiliar no tratamento de fisioterapia por meio do Kinect, o que torna-se uma prática bastante interessante e atrativa aos pacientes, possibilitando que eles próprios façam os seus exercícios, desde que o sistema já esteja funcionando corretamente e já esteja aprovado por fisioterapeuta habilitado.

O fisiogame desenvolvido funciona da seguinte maneira: ao abrir no jogo, o paciente deve selecionar quais exercícios pretende executar naquela seção de fisioterapia, a partir das opções exibidas no menu principal da tela inicial, como mostrado na Figura 3.

Figura 3: Interface inicial do fisiogame.

A seguir, o paciente é direcionado para outra interface, onde contém várias opções de exercícios para reabilitação de várias partes do corpo humano. Nesta interface, também há a descrição do modo de execução de cada exercício selecionado pelo paciente, como pode ser observado nas Figuras 4(a) e 4(b).



(a) (b)

Figura 4: Interfaces de seleção de exercícios do fisiogame.

Após selecionar todos os exercícios desejados, o paciente deve retornar ao menu inicial, para iniciar o jogo.

Na versão mobile do fisiogame utilizando Realidade Virtual e Aumentada, o procedimento para uso dos fisioterapeutas e pacientes é semelhante à versão da aplicação com Kinect mencionada anteriormente. Porém, o reconhecimento facial e gestual ocorre por meio de uma biblioteca chamada Cam2Play que torna possível esta funcionalidade com o auxílio da câmera do dispositivo móvel.

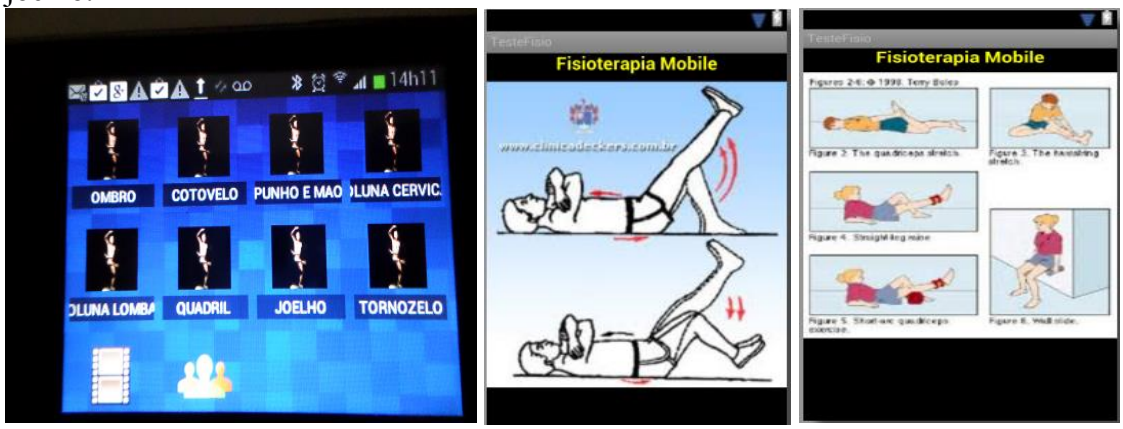
Em todas as linguagens de programação existem bibliotecas de funções que são muito úteis para o desenvolvimento de uma aplicação, pois economizam a escrita de rotinas e facilitam o desenvolvimento do software. As bibliotecas podem ser compartilhadas entre programas distintos, ou seja, permite que duas ou mais aplicações diferentes utilizem a mesma biblioteca ao mesmo tempo. Outra vantagem que se obtém ao elaborar uma biblioteca é que esta pode ser facilmente divulgada e disponibilidada por outros programadores.

A biblioteca Cam2Play foi desenvolvida utilizando o software Adobe Flash CS8, ou simplesmente Flash, que é uma plataforma de desenvolvimento gráfico que permite a criação de animações vetoriais em uma linha do tempo. A partir da versão 5.0 do Flash, foi incorporada a linguagem de programação ActionScript, orientado a objetos, esta tem fortes influências do JavaScript e permite criar aplicativos, aplicar filtros, gráficos em textos e imagens, criar animações, manipular objetos disponibilizados pelo Flash. Todos os recursos do Flash podem ser utilizados em navegadores que possuam o plugin especial do flash conhecido como Flash Player.

Na Figura 5(a), é apresentada a tela de seleção dos exercícios físicos no dispositivo móvel. Após escolher qual tipo de exercício físico se deseja fazer, são apresentadas telas que explicam como o exercício deve ser realizado pelo paciente.

Um fisioterapeuta habilitado deve pré-estabelecer os exercícios cadastrados no sistema, para que o paciente não faça-os de maneira incorreta. Nas Figuras 5(b) e 5(c) ,

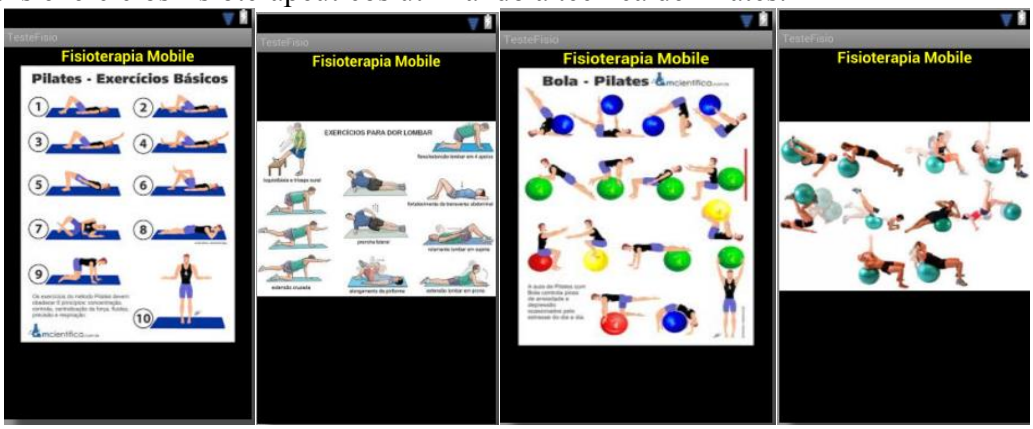
são exibidas algumas demonstrações de exercícios fisioterapêuticos realizados para o joelho.



(a) (b) (c)

Figura 5: (a) Tela de seleção de exercícios no celular; (b)(c) Exercícios de fisioterapia para joelho no celular.

Nas Figuras 6(a) e 6(b), são mostradas algumas explicações de exercícios fisioterapêuticos realizados para a coluna lombar. E nas Figuras 6(c) e 6(d) são exibidos alguns exercícios fisioterapêuticos utilizando a técnica de Pilates.



(a) (b) (c) (d)

Figura 6: (a)(b) Exercícios de fisioterapia para coluna no celular; (c)(d) Exercícios de fisioterapia utilizando a técnica de Pilates no celular.

Depois de selecionar os exercícios e ver as orientações para execução dos mesmos, o usuário deve iniciar a sessão de fisioterapia, de modo que a câmera do celular reconhecerá os movimentos do paciente, verificando se os exercícios estão sendo realizados de maneira correta. Na Figura 7, um paciente está realizando o seu cadastro de reconhecimento facial no dispositivo móvel para a realização dos exercícios de fisioterapia da aplicação.

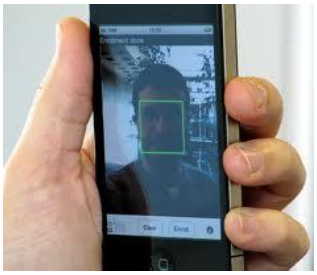


Figura 7: Reconhecimento facial do paciente no celular.

5. Conclusão

Em virtude do que foi mencionado, verifica-se que as inovações tecnológicas têm papel fundamental para o aperfeiçoamento, qualidade e diferenciação do tratamento fisioterapêutico de pacientes. Diante disso, o fisiogame utilizando o sensor Kinect e Realidade Virtual e Aumentada apresentado neste trabalho possibilita a manipulação dos dados dos pacientes imediatamente e posteriormente às seções de fisioterapia, com a finalidade de propiciar melhor análise e definição do plano de tratamento dos pacientes.

Desta maneira, percebe-se que esta aplicação é uma ferramenta de contribuição muito significativa para a área da saúde e medicina, uma vez que mostra um tratamento de fisioterapia mais personalizado e atrativo aos pacientes, desde que o sistema já esteja funcionando corretamente e já esteja aprovado por fisioterapeuta habilitado.

Logo, acredita-se na importância dos fisioterapeutas trabalharem com as novas tecnologias de RVA, pois proporcionam a visualização e interação do paciente com os exercícios fisioterapêuticos, facilitando o caminho para um tratamento mais harmonioso e eficaz.

Portanto, observa-se que a medicina é uma das áreas que mais demandam o uso de RVA em educação, 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.

6. Referências

Azuma, R. T.; Baiollot, Y.; Behringer, R.; Feiner, S.; Julier, S.; Macintyre, B. Recent advances in augmented reality. In: IEEE Computer Graphics and Applications, p. 34 47, 2001.

Birck, Fernando. Guia Prático para Iniciantes – Microsoft® XNA. Curitiba: Universidade UFPR, 2007.

Cardoso, A. Conceitos de realidade virtual e aumentada. [S.l.], 2007.

Costa, R. M. e Ribeiro, M. W. “Aplicações de realidade virtual e aumentada”. Porto Alegre: SBC, 2009. 146 p.

Gomes, W. L., Kirner, C. “Desenvolvimento de Aplicações Educacionais na Medicina com Realidade Aumentada”. Bazar: Software e Conhecimento Livre, N. 1, p 13-20, Julho, 2006.

Harma, A. et al. “Techniques and applications of wearable augmented reality audio”. In: *Audio Engineering Society Convention Paper*, Amsterdam, Holanda, 2003.

Kirner, C.; Zorzal, E. R. Aplicações educacionais em ambientes colaborativos realidade aumentada. In: XVI SBIE2005 â Simpósio Brasileiro de Informática na Educação, 2005.

Lamounier, E.; Cardoso, A. “Realidade virtual: uma abordagem prática”. São Paulo: Mania de Livro, 2004. 326 p.

Sistema Contra Roubos em Caixas Eletrônicos por meio de Reconhecimento Facial utilizando Redes Neurais Artificiais

Wildrey Barbosa Guimarães[1], Flávia Gonçalves Fernandes[1], Mariana Cardoso Melo[1], Bruno Gabriel Gustavo Leonardo Zambolini Vicente[1], Stéfano Schwenck Borges Vale Vita[1]

[1]Instituto de Engenharia e Tecnologia – Universidade de Uberaba (UNIUBE)
CEP: 38.408-343 – Uberlândia – MG – Brasil

wildrey@uniube.edu.br, flavia.fernandes92@gmail.com,
mariana.melo@uniube.br, bruno.vicente@uniube.br,
stefanoborges@gmail.com

Abstract. *The use of artificial intelligence to meet the growing demand of digital image processing is continuously evolving. During the last years, there has been a significant increase in the level of interest in mathematical morphology, artificial neural networks (ANN), in addition to processing, compression, and recognition images analysis systems based on knowledge. In this perspective, this paper presents the development of a system against thefts at ATMs that makes use of artificial neural networks technology. The proposed solution allows the installation of the system in all bank branches or in environments where such technology is needed, promoting the reduction of criminality.*

Resumo. *A utilização da inteligência artificial para suprir a crescente demanda do processamento de imagens digitais está evoluindo continuamente. Durante os últimos anos, tem havido um aumento significativo no nível de interesse em morfologia matemática, redes neurais artificiais (RNAs), além do processamento, compressão e reconhecimento de imagens em sistemas de análise baseados em conhecimento. Nesta perspectiva, este trabalho apresenta o desenvolvimento de um sistema contra roubos em caixas eletrônicos que faz uso desta tecnologia de redes neurais artificiais. A solução proposta permite a instalação do sistema em todas as agências bancárias ou em ambientes onde seja necessário tal tecnologia, promovendo a redução da criminalidade.*

1. Introdução

O reconhecimento facial é um dos atos que está em vigor no dia a dia, onde as pessoas precisam assegurar sua autenticidade, como na realização de transações bancárias, identificação em empresas, aeroportos, shopping centers entre outros. Os meios mais utilizados de identificação é através de senhas e uso de cartões com chips ou dispositivos magnéticos, onde o grande problema com esses meios é a facilidade de terceiros em obter as informações do usuário [Abate, 2007].

Esta atividade executada com tanta naturalidade por seres vivos, tem estimulado o interesse de pesquisadores que trabalham com a Inteligência Artificial.

O intuito das pesquisas sobre reconhecimento facial é desenvolver equipamentos que sejam capazes de realizar tal tarefa a fim de empregar essa tecnologia nas mais diversas atividades, como sistemas de vigilância, controles de acesso, entre outros. É uma área de pesquisa bem abrangente e envolve várias disciplinas como processamento de imagens, reconhecimento de padrões, visão computacional e redes neurais [Barret, 1998].

As redes neurais artificiais são um modelo computacional inspirado no cérebro humano e possuem a capacidade de aprender e generalizar através de exemplos. As redes neurais têm contribuído muito no desenvolvimento de sistemas de reconhecimento e classificação de padrões e são utilizadas em vários trabalhos voltados ao reconhecimento de expressões faciais [Azevedo, 2000].

Logo, o objetivo deste trabalho é a implementação de um sistema contra roubos em caixas eletrônicos, o qual é um software que atua na identificação facial e no acionamento de equipamentos, como o Gerador de Neblina e um sistema de alarme com monitoramento 24 horas, que ao receber um sinal de ativação, o aparelho automaticamente gera uma nuvem de vapor com aspecto de fumaça que preenche completamente o recinto protegido em menos de 30 segundos, reduzindo a visibilidade a menos de 30 cm, em todas as direções e também aciona outros dispositivos, como sistemas de alarme com monitoramento 24 horas, solicitando um atendimento emergencial.

2. Fundamentos do Reconhecimento Facial por meio de RNAs

As redes neurais artificiais têm contribuído muito no desenvolvimento de sistemas de reconhecimento de imagens. Existem diversos fatores que influenciam no reconhecimento facial humano tais como variação da iluminação, pose, expressão facial, óculos, mudanças nos cabelos, barba ou bigode [Faria, 2013].

O reconhecimento facial vem recebendo uma atenção significativa, principalmente durante os últimos anos, pode-se citar ao menos dois motivos para isto, onde o primeiro é o grande número e variedades de aplicações possíveis, sejam elas comerciais, militares ou de segurança pública e o segundo é a disponibilidade de tecnologias viáveis após décadas de pesquisa [Braga, 2000].

O reconhecimento facial é considerada uma área de pesquisa promissora. Existem diversas aplicações onde a identificação humana é necessária e o reconhecimento facial possui algumas vantagens sobre as demais tecnologias biométricas, pois é natural, não-intrusiva e de fácil utilização. A detecção de face consiste na utilização de métodos computacionais que verificam a existência de uma face, em uma determinada imagem digital, de vídeo ou fotografia [Haykin, 2001].

Um sistema de reconhecimento facial geralmente consiste de quatro módulos: detecção facial, normalização, extração de características e comparação. A detecção facial identifica a área referente a uma face em uma imagem e a isola do restante da imagem [Manssour, 2014].

A arquitetura de uma rede neural fica a critério do projetista, a qual deve ser testada para satisfazer seus objetivos. As RNA's de uma única camada são limitadas para resolver a grande parte dos problemas envolvendo reconhecimento facial [Gonzalez, 2000].

Para a obtenção de uma resposta mais satisfatória utilizaremos o Perceptron de Múltiplas Camadas (MLP – Multilayer Perceptron), de acordo com a Figura 1, que é uma rede com uma camada sensorial ou camada de entrada, que possui tantos nós de entrada quantos forem os sinais de entrada, uma ou mais camadas ocultas de neurônios e uma camada de saída com um número de neurônios igual ao número de sinais de saída, onde o sinal de entrada se propaga para frente através das camadas até a camada de saída, ou seja, é uma rede alimentada para frente [Carlos Junior, 2014].

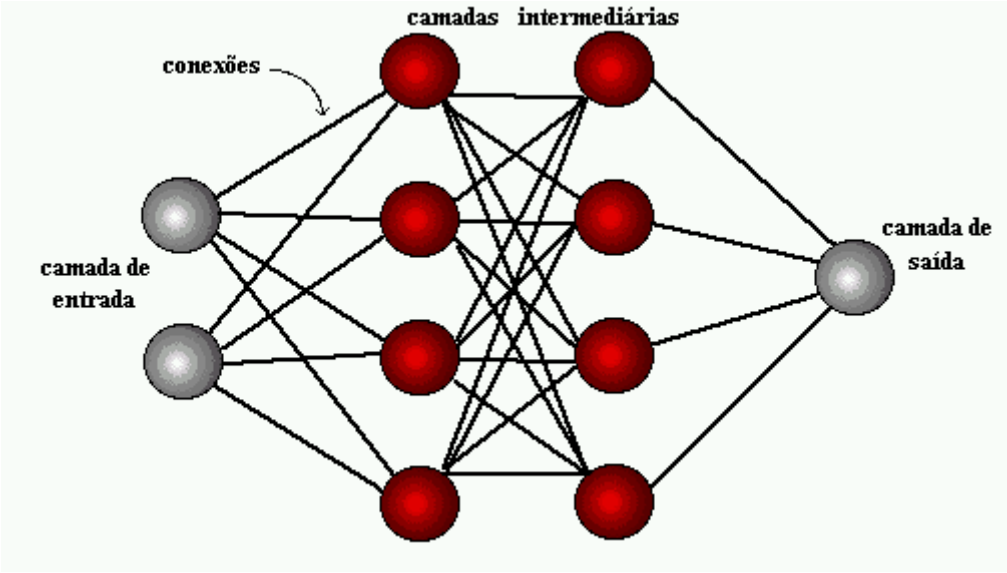


Figura 1: Estrutura da Rede Neural – MLP [Carlos Junior, 2014].

O uso de uma rede MLP com o algoritmo de retro propagação com muitos neurônios escondidos faz com que a solução seja atingida rapidamente durante o treinamento, mas o poder de generalização é sacrificado, ou seja, o desempenho em casos não vistos tende a piorar. Durante o treinamento, o sistema posiciona as funções discriminantes que classificam corretamente a maioria dos exemplos, para então lentamente classificar áreas com poucos exemplos. Por outro lado, o erro estabilizará em um alto valor se os graus de liberdade não forem suficientes [Deitel, 2011].

Um MLP com duas ou mais camadas escondidas é um aproximador universal, ou seja, realiza qualquer mapeamento entrada-saída. Isto acontece porque cada neurônio na primeira camada cria uma saliência e a segunda camada combina estas saliências em regiões disjuntas do espaço. Entretanto, um problema ainda sem solução é a determinação do número ótimo de camadas escondidas e do número de neurônios em cada camada para um dado problema [Carlos Junior, 2014].

Um classificador ótimo deve criar funções de discriminação arbitrárias que separem os agrupamentos de dados de acordo com a probabilidade a posteriori. A rede MLP pode fazer isto desde que: a) hajam neurônios suficientes para fazer o mapeamento; b) hajam dados em quantidade e qualidade; c) a aprendizagem convirja para o mínimo global; e d) as saídas estejam entre 0 e 1 com a soma das mesmas com valor unitário (neurônio softmax) [Azevedo, 2000].

Há procedimentos sistemáticos na busca pelo conjunto de pesos que produz um resultado desejado. Entretanto, esta busca deve ser controlada heuristicamente. O usuário atua na busca através da definição de quatro elementos: a) da seleção dos pesos iniciais;

b) das taxas de aprendizagem; c) do algoritmo de busca; e d) do critério de parada. Conjuntos finais de pesos diferentes surgem com mesma topologia e mesmo treinamento. Isto se deve a vários fatores, entre eles o fato de existirem muitas simetrias no mapeamento entrada-saída, a não existência de garantias que o problema tenha somente uma solução, e às condições iniciais aleatórias do conjunto de pesos. Deve-se sempre lembrar que a aprendizagem é um processo estocástico, e, portanto, deve-se treinar cada rede várias vezes com condições iniciais diferentes e usar a melhor [Haykin, 2001].

Em redes não-lineares, como praticamente sempre é o caso de uma rede MLP, a seleção do tamanho do passo é muito importante. Em princípio deseja-se que o passo seja o maior possível no início do treinamento para acelerar a convergência, mas ao longo do processo este valor deve diminuir para se obter uma maior precisão do resultado [Braga, 2000].

3. Metodologia

Em relação aos aspectos metodológicos e tecnológicos, foi utilizada a rede neural MLP para a implementação do sistema, onde ela é submetida ao treinamento com o algoritmo de Perceptron. A rede consiste de um conjunto de unidades sensoriais (nós de fonte) que constituem a camada de entrada, uma ou mais camadas ocultas de nós computacionais e uma camada por camada. Estas redes neurais são normalmente chamadas de *perceptrons de múltiplas camadas* (MLP, multilayer perceptron) [Deitel, 2011].

Desse modo, a rede neural artificial é treinada, onde recebe imagens do banco de dados para executar esse treinamento. As redes neurais tem como principal característica a capacidade de aprendizado através de exemplos, pois é esta característica que difere a abordagem conexionista da inteligência artificial simbólica. A aprendizagem é um processo pelo qual os parâmetros livres de uma rede neural são adequados através de um processo de estimulação pelo ambiente no qual a rede está inserida [Ludwig & Costa, 2007].

Além disso, foi utilizada linguagem de programação Java, framework Encog, que é uma estrutura avançada de aprendizado de máquina que suporta uma variedade de algoritmos avançados, bem como aulas de apoio para normalizar e processar dados. O Encog Framework está em desenvolvimento ativo desde 2008 e está disponível para Java, .Net e C / C + +.

Neste trabalho, o banco de imagens escolhido para testes foi o ORL que foi produzido pela Olivetti Research Laboratory em Cambridge, UK. Este banco é gratuito/público e possui um total de 400 imagens sendo 40 indivíduos e 10 imagens diferentes para cada indivíduo. As imagens possuem variações na expressão facial (olhos abertos/fechados, sorrindo/sem sorrir), iluminação e detalhes faciais (com ou sem óculos). As imagens foram obtidas sob um fundo escuro e homogêneo, e estão em escala de cinza com uma resolução de 92x112 pixels [AT&T, 1992].

Este banco de dados foi escolhido por ser público, conter uma quantidade considerável de imagens, possuir as variações necessárias para o experimento (iluminação, pose, expressão facial), outro motivo é o fato deste banco de dados ser bastante utilizado em trabalhos na área possibilitando uma comparação de resultados com trabalhos correlatos.

A câmera utilizada para desenvolvimento do protótipo foi uma micro câmera colorida com áudio, conforme exibido na Figura 2.

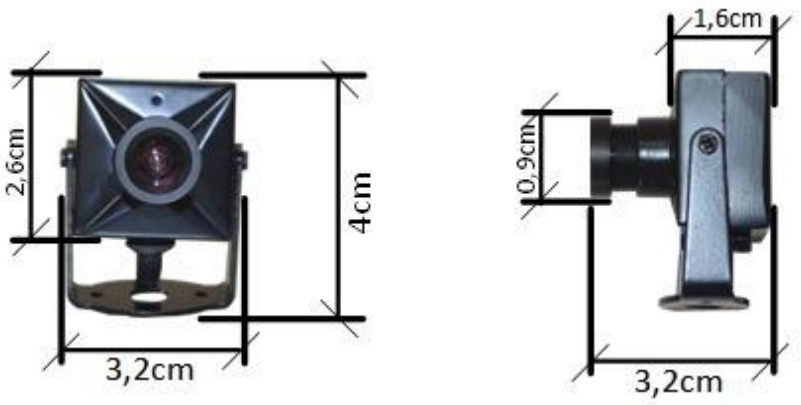


Figura 2: Micro câmera 420 linhas com áudio.

Suas especificações técnicas são as seguintes:

- Sensor de imagem – CMOS ¼";
- Sistema de vídeo – NTSC/PAL;
- Lente de 3,6 mm;
- Pixels efetivos – 628 (H) X 582 (V);
- Resolução horizontal de 420 linhas;
- Iluminação mínima de 0,2 lux;
- Alimentação de + 9Vdc ~ + 13,5Vdc;
- Consumo de 0,2W;
- Conector de vídeo RCA.

O microcontrolador aplicado é um ATmega328, que pode ser visualizado na Figura 3, o qual possui 14 pinos de entradas/saídas digitais (dos quais 6 podem ser usados como saídas PWM), 6 entradas analógicas, 4 UARTs (portas seriais de hardware), um oscilador de cristal de 16 MHz, uma conexão USB, uma entrada de alimentação e um botão de reset.

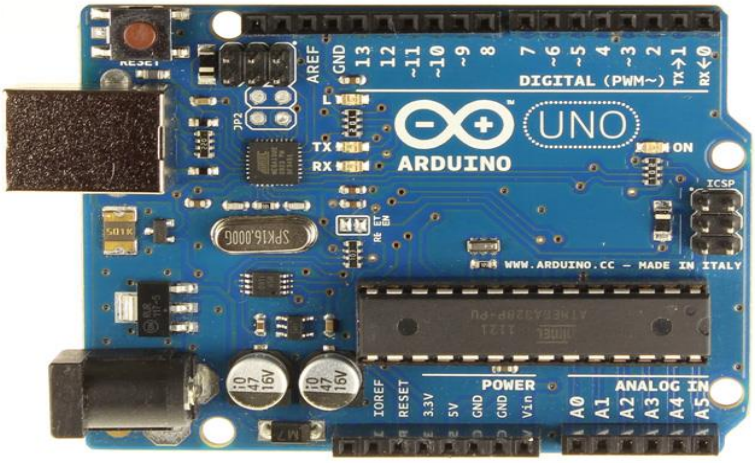


Figura 3: Arduino Uno.

3. Resultados e Discussões

O protótipo do sistema contra roubos em caixas eletrônicos por meio de reconhecimento facial utilizando redes neurais artificiais consiste em treinar a rede a partir de fotografias da face de indivíduos. Após treinada, a rede é submetida ao teste de reconhecimento onde recebe uma outra imagem de um dos indivíduos com o qual a rede foi treinada anteriormente.

Neste trabalho, o banco de imagens foi implementado através da aquisição de fotos de cada cliente de uma agência bancária, onde essas fotos são analisadas pelo sistema no momento em que o cliente adentrar na agência para efetuar alguma operação, após o horário de expediente.

A construção do protótipo foi dividida em módulos, que integrados formam o dispositivo contra roubo em caixas eletrônicos. Cada módulo representa um conjunto de funcionalidades que são essenciais ao projeto, para cumprir com os objetivos iniciais. O fluxograma abaixo mostra todos os módulos de maneira integrada. A Figura 4 mostra o fluxograma geral da implementação do sistema.

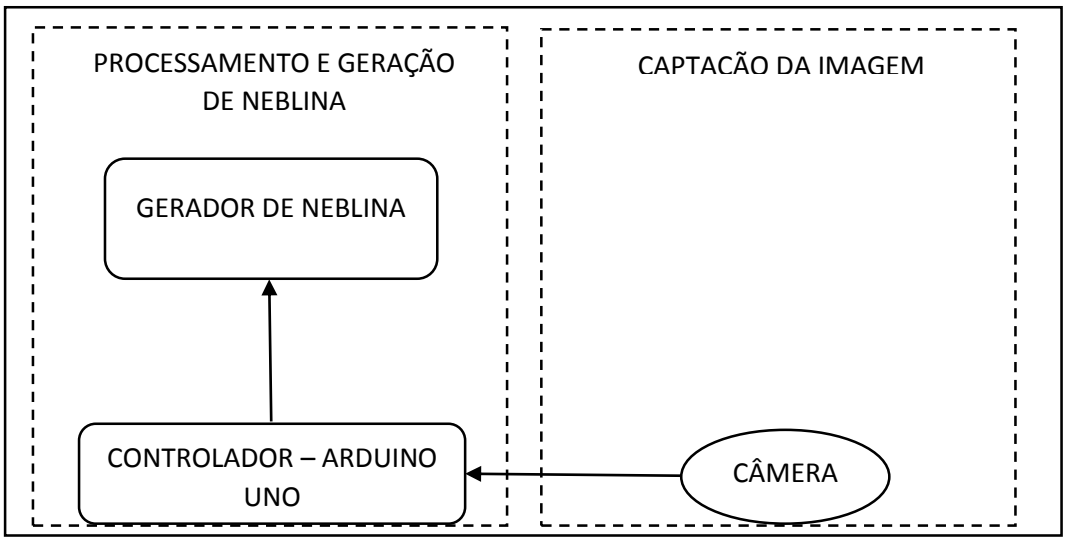


Figura 4: Fluxograma geral do protótipo.

Enquanto isso, a estrutura física do sistema foi desenvolvida utilizando o software AutoCad, onde no interior da estrutura está o controlador arduino UNO. Utilizou-se como material para a construção da estrutura placas de acrílico visando uma fácil montagem e também para ter uma visibilidade do microcontrolador para futuras manutenções. Desenvolveu-se a modelagem 3D de uma estrutura compatível com os requisitos do projeto.

O gerador de neblina, como pode ser visto na Figura 5, é uma evolução no suprimento de segurança das agências, que ao receber um sinal de ativação de um sensor próprio ou externo, o aparelho, automaticamente, gera uma nuvem de vapor com aspecto de fumaça, de acordo com a Figura 6, que preenche completamente o recinto protegido em menos de 30 segundos e reduz a visibilidade a menos de 30 cm, em todas as direções. Assim, não há como fazer qualquer deslocamento dentro do ambiente protegido, inibindo totalmente a ação de possíveis delitos.

Figura 5: Gerador de neblina.

Figura 6: Momento em que o gerador de neblina é acionado.

A construção do protótipo foi concluída, resultando em uma estrutura escalar pequena, porém com placa de controle flexível, que suporta o acionamento do gerador de neblina.

4. Conclusão

Em virtude do que foi mencionado neste trabalho, o desenvolvimento do protótipo foi concluído conforme os objetivos iniciais, de maneira que é uma ferramenta muito útil no combate à criminalidade, principalmente roubos em caixas eletrônicos, o que propicia maior segurança aos estabelecimentos que possuem este novo dispositivo implementado.

As experiências acumuladas culminaram na geração de sólidos conhecimentos relacionados a redes neurais artificiais, microcontroladores, desenvolvimento de projetos e linguagem de programação. Com isso, foi validado os conhecimentos adquiridos, aplicando-os no problema de reconhecimento facial e implementando um sistema capaz de reconhecer a face humana.

Conclui-se, por conseguinte, com base nos resultados obtidos na programação deste projeto que as metas foram alcançadas e superadas, abrindo caminho para novos cenários. Além disso, foi possível demonstrar a eficiência de uma rede do tipo MLP aplicada na resolução do problema do reconhecimento facial.

Como trabalhos futuros, pretende-se aperfeiçoar e melhorar o funcionamento sistema, principalmente no que se refere ao reconhecimento facial de imagens, a fim de que o mesmo seja utilizado com bom desempenho e eficácia em estabelecimentos

comerciais interessados nesta tecnologia, além de inserir o maior número de imagens de clientes possível no banco de dados do protótipo.

5. Referências

Abate, A. et al. 2D and 3D face recognition: A survey. Pattern Recognition Letters, Elsevier Science Inc., New York, NY, USA, v. 28, n. 14, p. 1885–1906, out. 2007. ISSN 01678655. Disponível em: <http://dx.doi.org/10.1016/j.patrec.2006.12.018>.

AT&T. The Database of Faces, Cambridge University Computer Laboratory. Disponível em: <http://www.cl.cam.ac.uk/ research/dtg/attarchive/facedatabase.html>. Acesso em 10 Maio 2011. 1992. Disponível em: <http://www.cl.cam.ac.uk/research/dtg/attarchive/facedatabase.html>.

Azevedo, F. M.; BRASIL, L. M.; OLIVEIRA, R. C. L. Redes Neurais com Aplicações em Controle e em Sistemas Especialistas. [S.l.]: VISUAL BOOKS, 2000. ISBN 8575020056.

Barret, W. A. A survey of face recognition algorithms and testing results. ACM Transactions on Computer Systems, 1998.

Braga, A.; Carvalho, A. C.; Ludermir, T. B. Redes Neurais Artificiais: Teoria e aplicações. [S.l.]: LTC Editora, 2000.

Carlos Junior, Luis Fernando Martins. Reconhecimento facial utilizando redes neurais. Centro Universitário Eurípedes de Marília. Disponível em: <aberto.univem.edu.br/handle/11077/360:>. Acesso em: jan. 2014.

Deitel, P.; Deitel, H. Como programar C. 6 ed. São Paulo: Pearson Prentice Hall, 2011. 818 p.

Faria, Alessandro O. Biometria: reconhecimento facial livre. Disponível em <http://www.linhadecodigo.com.br/artigo/1813/biometria-reconhecimento-facial-livre.aspx>. Acesso em: 18 dezembro 2013.

Gonzalez, Rafael C.; Woods, Richard E. Processamento de imagens digitais. São Paulo: Editora Blucher. 2000. 509 p.

Haykin, S. Redes neurais: princípios e prática. 2ª ed. Porto Alegre: Bookman, 2001. 900 p.

Ludwig Jr., O.; Costa, E. M. Redes neurais: fundamentos e aplicações com programas em C. Rio de Janeiro: Editora Ciência Moderna Ltda. 2007. 125 p.

Manssour, Isabel H.; Pinho, Marcio S. Manipulação de Imagens: programação de software básico. Disponível em: <http://www.inf.pucrs.br/~pinho/PRGSWB/Exercicios/AulaImagens/>. Acesso em: 02 janeiro 2014.

AVALIAÇÃO DO USO DE ASPECTOS SOCIAIS PARA ROTEAMENTO OPORTUNISTA CONSIDERANDO ZONAS RURAIS

Bryan Larry Pond[1], Antonio C. Oliveira Júnior[1], Rosario Ribeiro[1], Waldir Moreira[2]

[1]Universidade Federal de Goiás (UFG/DCC), Brasil

[2]Universidade Lusófona, COPELABS, Lisboa - Portugal

bryan.pond2012@gmail.com, antonio@catalao.ufg.br, rosariocamposribeiro@gmail.com, waldir.junior@ulusofona.pt

Abstract. *In this paper is addressed the use of social aspects (e.g. notion of communities and users' daily routines) to support routing in opportunistic networks. It is described concepts and routing protocols based in social aspects. Initial results show the viability of such protocols using the ONE simulator (Opportunistic Network Environment simulator).*

Resumo. *Neste artigo é abordado a utilização de aspectos sociais (e.g. noção de comunidades e rotinas diárias dos usuários) como suporte ao encaminhamento em redes oportunistas. Apresenta-se conceitos e protocolos de roteamento oportunista baseado em aspectos sociais. Resultados iniciais demonstram a viabilidade de tais protocolos utilizando o simulador ONE (Opportunistic Network Environment simulator).*

1. Introdução

Com a globalização das tecnologias de comunicação e o crescente aumento na demanda por comunicação surgem problemas básicos como a falta de infraestrutura para a prestação de tais serviços, os quais incentivaram tecnologias alternativas como as redes oportunistas.

Redes oportunistas [Moreira 2013] estão cada vez mais promissores para encaminhamento por necessidade, onde a informação transita por nós intermediários até o destino, estratégia denominada *store-carry-and-foward*. Nesse modelo em comparação à internet convencional a transmissão acontece oportunistamente, de acordo com necessidade e oportunidade de envio e de modo que a comunicação pode ocorrer mesmo que não exista uma rota que una diretamente os nós e sem a necessidade de conhecimento prévio da topologia.

Estratégias como redes oportunistas também são visadas como forma de inclusão digital, levando conectividade há zonas remotas onde não há infra-estrutura para o modelo convencional de Internet. Conectividade em zonas rurais, que foi o foco deste trabalho, vem recebendo especial atenção em projetos relacionados a redes oportunistas.

Neste artigo é abordado o fator social como fonte de informação e suporte, para a tomada de decisões de protocolos de roteamento em redes oportunistas em um cenário

rural próximo a cidade de catalão. Apresentam-se descrições de conceitos relacionados a redes oportunistas e protocolos de encaminhamento baseado em aspectos sociais (e.g. noção de comunidades e rotinas diárias dos usuários). Utilizando o simulador ONE (*Opportunistic Network Environment*), é apresentado um estudo comparativo entre protocolos que utilizam e não o fator social para encaminhamento, *dLife* [Moreira 2012] e *Epidemic* [Hui 2011] a fim de demonstrar as vantagens de abordagens sociais. Este artigo é organizado da seguinte forma. Após a introdução, a seção II descreve os conceitos de redes oportunistas juntamente com a arquitetura DTN, tipos de contatos e cenários de aplicações. A seção III dedica-se a protocolos de roteamento que utilizam aspectos sociais como suporte ao encaminhamento. Na seção IV é apresentado os resultados de performance de dois protocolos em um cenário real utilizando o simulador ONE. As conclusões deste trabalho são discutidas na seção V.

2. Redes Oportunistas e DTN

Redes oportunistas tem demonstrado ser soluções atrativas quando a Internet convencional não se mostra viável. Aplicações em ambientes onde não dispõe-se de uma infra-estrutura de comunicação adequada são bons exemplos de cenários que podem usufruir de uma estratégia oportunista.

O conceito de redes oportunistas segue a abordagem de redes tolerantes a atrasos e desconexões (DTNs) [Fall 2003], onde não há necessidade de conhecimento da topologia, e o encaminhamento é feito oportunistamente ao longo dos contatos esporádicos que ocorrem ao passar do tempo.

Entretanto, as redes oportunistas são altamente dinâmicas, composta por dispositivos móveis e fixos e tiram vantagem dos contatos oportunistamente ao longo da variação do tempo para trocar informações carregadas pelos usuários. Essas redes costumam ser compostas por dispositivos móveis com restrição de bateria, assim, eficiência energética tem recebido especial atenção nos últimos anos. Então, mecanismos para suporte ao roteamento energeticamente eficiente devem ser considerados [Junior 2012].

Originalmente o conceito inicial de DTNs foi motivado para melhorar a comunicação interplanetária da época [Fall 2003]. Posteriormente esse conceito passou para ambientes de rede além do espacial, propiciando assim o conceito de redes oportunistas. Para esses cenários de rede a arquitetura TCP/IP (usada na internet convencional) não é ideal.

Em ambientes onde a comunicação levaria minutos, horas ou dias se mostra inviável a utilização de tal arquitetura, desta forma surgiu a proposta de uma rede baseada no esquema de *store-carry-and-foward* (SCF). Para usar a estratégia de SCF temos uma nova camada na pilha de protocolas a *Bundle layer* (camada de agregação) funcionando abaixo da camada de aplicação.

A camada de agregação usa armazenamento persistente para armazenar mensagens, lidando com os problemas de atraso e conexão, que na internet seriam perdidas [Moreira 2012]. A comunicação funciona enviando a mensagem (*bundles* ou agregados) inteira ou em parcelas por nós intermediários até alcançar o destinatário. Se a

mensagem for dividida pela camada de agregação para ser enviada, ela é reconstruída na camada de agregação do destino.

Para transmitir nó a nó, usando o paradigma store-carry-and-foward o transmissor pode se valer da transferência de custódia [Fall 2003]. Quando um nó tem a custódia de um *bundle* (agregado), ele transmite esse agregado com um pedido de transferência de custódia e aguarda um *ack*. Se o nó receptor aceitar essa transferência de custódia ele retorna um *ack* para que o agregado possa ser apagado no nó anterior.

2.1. Tipos de contatos

Em redes com conexões intermitentes, em contrariedade a Internet convencional, temos a característica de que um nó nem sempre pode estar disponível, assim o conceito de contato é um importante fator de consideração. Um contato corresponde a uma ocasião favorável para os nós trocarem dados [Melo 2011]. No conceito de DTNs contatos podem ser classificados como persistentes, sob demanda, previsíveis, programados e oportunistas.

2.1.1. Contatos persistentes

São contatos que estarão sempre disponíveis. Uma conexão de usuário com a Internet é um exemplo de contato persistente.

2.1.2. Contatos sobre demanda

Precisam ser acionados e então se comportam como contatos persistentes. Um exemplo seria um usuário usando uma conexão discada de Internet [Melo 2011].

2.1.3. Contatos previsíveis

Quando os nós podem fazer previsões sobre o horário e duração do contato com algum nível de segurança, temos um contato previsível. Uma rede em uma área rural, exemplificada na Figura 1, onde um ônibus funciona como intermediário para transportar a informação seria um bom exemplo de contato previsível, baseado em histórico teríamos o horário que o ônibus chega ao ponto e o tempo que fica parado. Porém previsões estão sujeitas a erros, visto que não se pode prever acidentes ou engarrafamentos no caso do ônibus por exemplo.

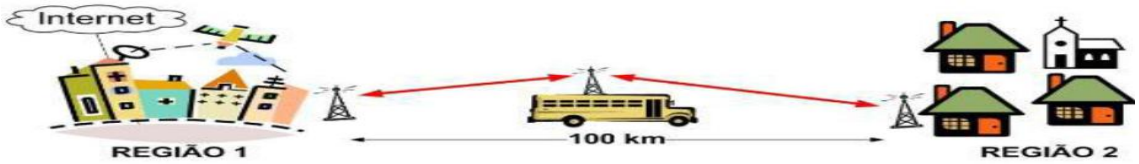


Figura 1. Contato previsível de uma DTN em zona rural [Melo 2011].

2.1.4. Contatos programados

Pode-se estabelecer uma rotina de comunicação entre nós, quando ocorrerá o contato e a sua duração, chamemos de agenda de contato. Para contatos programados é necessária uma sincronização entre os nós participantes.

2.1.5. Contatos oportunistas

São aqueles que não se pode prever ou definir momento nem duração do contato. Não há nenhuma conexão direta ou conhecimento da topologia pelos nós, são contatos

puramente aleatórios entre os nós da rede até alcançar o destinatário. Um exemplo seria uma *Pocket-switch-network* [Hui 2005] formada pelos aparelhos de usuários em uma universidade por exemplo.

2.2. Cenários

Como a topologia da rede pode estar sempre mudando, nós estão em movimento ou nem sempre estão disponíveis, assim surge a necessidade de protocolos específicos para lidar com a variabilidade da topologia.

Na literatura o grau de conhecimento sobre a topografia da rede e suas variações costuma ser dividido em dois cenários, o cenário dinâmico e o determinístico, cada qual com várias propostas de roteamento.

2.2.1. Cenário dinâmico

Nesse cenário a movimentação dos nós não é completamente conhecida e as rotas não podem ser calculadas. Assim os nós não tem conhecimento prévio do estado da rede, os protocolos para esse tipo de cenário sugerem que oportunistamente a informação seja transmitida para outro nó até alcançar o destinatário.

A proposta mais simples de roteamento para cenários dinâmicos é o roteamento epidêmico [Vahdat 2000]. Nessa proposta de roteamento como é característico de cenários dinâmicos, pressupõe-se que o nó de origem não tem nenhuma informação quanto à topologia da rede. A estratégia de transmissão usada é semelhante a uma epidemia onde o nó fonte transmite para todos os nós que venha a encontrar até que a mensagem consiga atingir seu destino. Ou seja, quanto maior o número de nós atingidos maior a chance de a mensagem ser entregue ao destinatário. Contudo essa estratégia apresenta múltiplas desvantagens como desperdício de recursos.

2.2.2. Cenário determinístico

Temos um cenário determinístico quando os contatos, sua duração e a topologia da rede são conhecidos pelos nós. Com isso os nós podem conhecer os contatos e o caminho até o destino.

Um dos modelos propostos para cenários determinísticos é o modelo de grafos evolutivos [Shingo Mabu 2007]. Esse modelo baseia-se no conceito de que se os nós da rede conhecerem o momento em que determinado enlace estiver ativo, ele pode escolher o melhor caminho para enviar a mensagem até o destino.

3. Protocolos de roteamento oportunista com base social

Como fazer o encaminhamento de mensagens em redes oportunistas é um fator crucial para o desempenho da mesma. Atualmente o uso de redes sociais (como facebook e twitter) alcançou uma escala global, associado a isso o uso de aparelhos de comunicação móveis, muitas vezes usados para acessar alguma rede social, também aumentou significativamente, devido a fatores como menor custo e maior facilidade de acesso á tecnologia comparado com alguns anos atrás. Segundo relatório da Cisco sobre previsão do crescimento do tráfego móvel global, até 2015 haverá um smartphone por habitante no planeta.

Tratando-se de encaminhamento oportunista, podemos tirar vantagem da interação social e do aumento no uso de dispositivos móveis para criar ou melhorar protocolos de roteamento em redes oportunistas através de aspectos sociais [Moreira 2012].

Já foram propostos vários protocolos que podem variar de acordo com o encaminhamento empregado. Usar algum teor social como fator de decisão no roteamento oportunista tem grandes vantagens na entrega da informação de forma eficiente. A estratégia com caráter social avaliada nesse artigo foi o dLife [Moreira 2012].

3.1. dLIFE

O dLIFE (Opportunistic Routing based on Users Daily Life routines) [Moreira 2013] funciona na rotina diária dos nós e considera duas funções de utilidade complementares. A primeira o Time-Evolving Contact Duration (TECD), que determina o peso social entre os usuários baseado em sua interação social durante suas rotinas diárias. E Time-Evolving Importance(TECDi) que mensura a importância do nó considerando seus vizinhos e respectivos pesos sociais.

Estruturas sociais são compostas pelos usuários (pelos nós), e podem mudar constantemente. Quando um nó conhece outro sua rede pessoal muda e conseqüentemente toda a estrutura social a qual aquele nó pertence também muda. Considerando isso o dLife, com a TECD consegue captar o dinamismo do comportamento social dos usuários, captando o funcionamento de suas rotinas de modo mais eficiente do que uma estimativa por histórico que é a estratégia mais usada.

Assim, o encaminhamento realizado pelo dLife considera o peso social do nó que carrega a mensagem em relação ao destino bem como o peso social do nó intermediário a este mesmo destino. Nos casos onde o nó intermediário tem um peso social (i.e. tem uma relação social forte com o destino, o nó fonte envia uma copia da mensagem ao nó intermediário. Caso contrario, a importância (TECDi) do nó que carrega a mensagem e do no intermediário é levada em conta na hora de replicar a informação. Isto é, o no intermediário recebera uma copia no caso de ser mais importante que o nó que detém a informação naquele momento.

4. Avaliação de desempenho

Esta seção dedica-se a uma simulação demonstrando o uso do protocolo dLife. Essa analise foi feita através do simulador de redes ONE (Opportunistic Network Enviroment) [Moreira 2012].

Os resultados de performance de protocolos oportunistas costumam ser avaliados seguindo as principais métricas de performance que são: probabilidade de entrega (diferença entre o número de mensagens entregues e o total de mensagem criadas), custo (numero de mensagens replicadas por cada mensagem criada), e o atraso (tempo entre a criação e entrega da mensagem).

4.1. Cenário avaliado

A simulação foi realizada usando um cenário, mostrado na Figura 2, das cidades e área rural no entorno da linha ferroviária que passa pela cidade de Catalão, com nós Wi-Fi 802.11 distribuídos em grupos. Os grupos de veículos seguem o movimento a uma velocidade de 50 a 110 km/h. Os trens tem velocidade de 60 a 80 km/h e faz a rota entre as cidades da região cerca de 60km de catalão até Ipameri. Os barcos se limitam a movimento na hidrovia e se movem de 20 a 55 km/h.

Estão posicionadas na região 13 fazendas estáticas na região que abrangem a área de cobertura do trilho do trem e da rodovia de acordo com o padrão Wi-Fi usado. A área de cobertura ao redor dos nós é de 100 metros em área aberta.

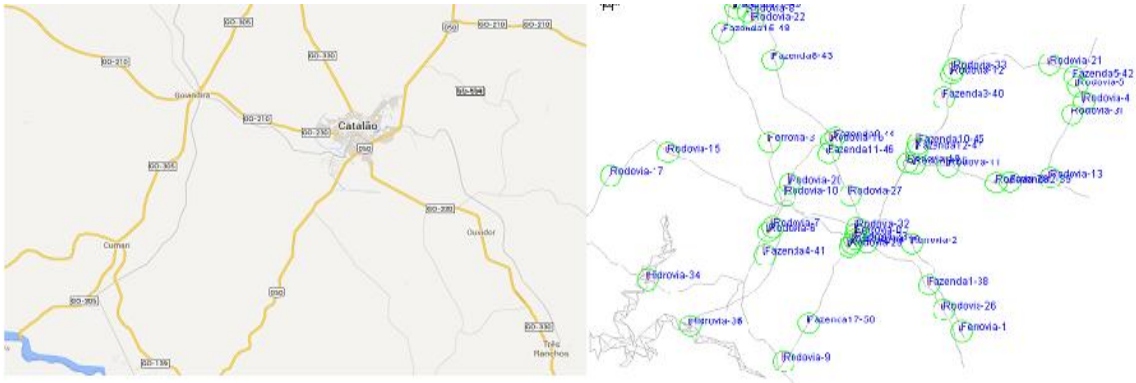


Figura 2. Mapa do cenário e Simulação rodando no cenário escolhido com o roteamento epidêmico.

Os valor de TTL das mensagens é de 1 dia. As mensagens tem tamanho de 2 MB. E o buffer é de 512 MB.

4.2. Resultados

Os resultados ainda superficiais tem por objetivo uma comparação entre os protocolos considerando o TTL de acordo com as métricas de performance.

A figura 3 mostra a probabilidade de entrega dos dois protocolos no cenário avaliado, com TTL 1 dia. Podemos ver que o dLIFE apresenta um probabilidade de entrega superior devido ao grande desperdício de recursos com replicação do epidemic gerando muitas mensagens e diminuindo a taxa de replicas entregues.

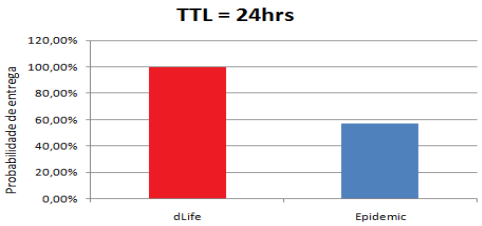


Figura 3. Probabilidade de entrega com TTL de 1 dia.

A Figura 4 apresenta o atraso de entrega em minutos de com TTL de 1 dia. O protocolo dLife claramente apresenta um atraso em torno de menor que o roteamento epidêmico. Pelo modo como o dLIFE toma decisões de envio (independente da noção de comunidade) quando há fortes laços sociais com o destino ou encontros rotineiros para aumentar a probabilidade de entrega da mensagem. Como a rede do cenário tem

bastante interação entre os nós o algoritmo rapidamente se adapta a "comunidade" do cenário, enquanto o roteamento epidêmico acaba criando muitas réplicas, que levam muito tempo para chegar ao destino.

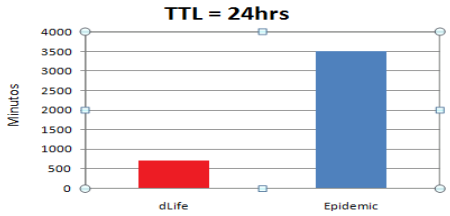


Figura 4. Atraso (minutos) com TTL de 1 dia.

A Figura 5 apresenta a quantidade de réplicas por mensagem, o custo, com TTL de 1 dia. Os resultados se devem ao protocolo dLife ter visão do grafo social e seus nós mais importantes em qualquer instante, independe da noção de uma comunidade diminuindo assim a necessidade de réplicas. Já o epidemic como é natural de sua abordagem gera uma maior taxa de replicação.

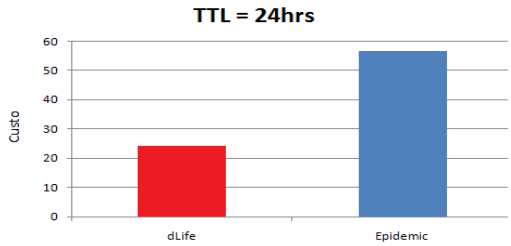


Figura 5. Custo (réplicas) com TTL de 1 dia.

5. Conclusões

Este trabalho visou um estudo de redes oportunistas como alternativa quando o modelo de Internet convencional não se mostra a melhor opção. Nesse conceito abordou-se fatores sociais como suporte ao encaminhamento em protocolos de roteamento oportunistas.

Após conceitos relacionados a redes oportunistas, aplicações e protocolos de roteamento, foi mostrado resultados de uma breve comparação entre protocolos a fim de demonstrar as vantagens de protocolos com base social em opção a abordagens mais rústicas.

Como trabalhos futuros, pretende-se avaliar exaustivamente todos os protocolos de roteamento oportunista com base em aspectos sociais considerando também outros cenários e métodos de simulação. Para além de simulações, pretende-se propor e avaliar os protocolos em um ambiente de teste real.

Agradecimentos

Este trabalho é financiado pela Fundação de Amparo à Pesquisa do Estado de Goiás (FAPEG) através do projeto número 201200544420886 (Edital 005/2012 - Universal).

References

A. Junior, R. Sofia, A. Costa, Energy-awareness in multihop routing," IFIP Wireless Days (WD), pp. 1–6, November, 2012.

A. Keranen, J. Ott, and T. Karkkainen, "The ONE Simulator for DTN Protocol Evaluation," in SIMUTools'09, (Rome, Italy), March 2009.

A. Vahdat, D. Becker. "Epidemic Routing for Partially Connected Ad Hoc Networks" (2000).

Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Up-date, 2011-2016. Relatório técnico.

K. Fall, "A delay-tolerant network architecture for challenged Internets", ACM SIGCOMM, 2003.

M. Melo, "Proposta e Avaliação da Utilização do Tráfego Aeroviário Nacional como Uma Rede Tolerante a Atrasos e Desconexões" Dissertação de Mestrado. 2011.

Moreira, W.; Mendes, P, " Social-Aware Opportunistic Routing: The New Trend," Routing in Opportunistic Networks, Springer New York, pp.27-68, 2013.

Moreira, W.; Mendes, P.; Sargento, S., "Opportunistic routing based on daily routines," World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2012 IEEE International Symposium on a , vol., no., pp.1,6, 25-28 June 2012.

Moreira W., Mendes P., Ferreira R., Cirqueira D., Cerqueira E., "Opportunistic Routing based on Users Daily Life Routine" "draft-moreira-dlife-02" Internet-Draft.

Mtibaa, A.; May, M.; Diot, C.; Ammar, M., "PeopleRank: Social Opportunistic Forwarding," INFOCOM, 2010 Proceedings IEEE , vol., no., pp.1,5, 14-19 March 2010.

Ott, J.; Kutscher, D., "A disconnection-tolerant transport for drive-thru Internet environments," INFOCOM 2005 24th Annual Joint Conference of the IEEE Computer and Communications Societies, vol.3, no., pp.1849,1862 vol. 3, 13-17 March 2005.

Pan Hui; Crowcroft, J.; Yoneki, E., "BUBBLE Rap: Social-Based Forwarding in Delay-Tolerant Networks," Mobile Computing, IEEE Transactions on , vol.10, no.11, pp.1576,1589, Nov. 2011.

Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. 2005. "Pocket switched networks and human mobility in conference environments". In Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking (WDTN '05).

Shingo Mabu, Kotaro Hirasawa, and Jinglu Hu. "A Graph-Based Evolutionary Algorithm: Genetic Network Programming (GNP) and Its Extension Using Reinforcement Learning". Evol. Comput. 15, 3 (September 2007), 369-398.

Ting Liu, Christopher Sadler, Pei Zhang, and Margaret Martonosi. "Implementing Software on Resource-Constrained Mobile Sensors: Experiences with Impala and ZebraNet". To appear in the Second International Conference on Mobile Systems, Applications, and Services (MobiSys 2004), June 2004.

Construção de um Processador de Linguagem Natural

Iohan Gonçalves Vargas[1], **Vaston Gonçalves da Costa**[1]

[1]Universidade Federal de Goiás - Campus Avançado de Catalão (UFG - CAC)
Catalão – GO – Brasil – Departamento de Ciência da Computação

{iohanufg,vaston}@gmail.com

Abstract. *Natural language processors emerge as keys tools for formal verification of systems. The use of this type of translator supplies a lack in proper validation formulas and their application cost is reasonable. Through a Natural Language Processor, which helps a user to process a text in natural language into logical formulas in DL, whose syntax is formal, tough learning and mastery, can construct a knowledge base and then make inferences about the information in order to explore the database through a logical reasoner. Thus, this paper presents how one can use translators with logical reasoners in the presentation of knowledge in a specific domain and generate and explain structured on the same evidence.*

Resumo. *Processadores de linguagem natural surgem como principais ferramentas para verificação formal de sistemas. A utilização desse tipo de tradutor supre a carência na validação correta de fórmulas e seu custo de aplicação é acessível. Dado um Processador de Linguagem Natural, que auxilia a processar um texto, em linguagem natural, para fórmulas lógicas em DL (Description Logic), cuja sintaxe é formal, de difícil aprendizado e domínio, pode-se construir uma base de conhecimento e, posteriormente, realizar inferências sobre as informações. Assim, este trabalho visa apresentar como se podem utilizar tradutores lógicos na representação do conhecimento em um domínio específico de forma que possibilite raciocinadores tomarem decisões.*

1. Introdução

Com os avanços teóricos, representação do conhecimento surge como um dos conceitos centrais e importantes em Inteligência Artificial (IA). Segundo Davis [1993], é um ambiente pelo qual se modela dados, de forma a orientar o raciocínio sobre um determinado domínio, objetivando facilitar a tomada de decisões, a partir de inferências, tendo como base a representatividade real. A definição de *Description Logic* - DL (Lógica Descritiva) tem sido amplamente utilizada em Bases de Conhecimento, é um conceito capaz de compreender diferentes formas, na qual pode se representar o conhecimento e obter conclusões. Em meados da década de 70, surgiu o conceito de DL, que se define basicamente em um nome, que se refere a qualquer uma das várias linguagens lógicas comumente utilizadas em Representação do Conhecimento. Assim, define-se um conjunto de conceitos, tidos como afirmações individuais (ABOX) e um conjunto de relações binárias definidas sobre os mesmos (TBOX) [Davis *et al.*, 1993].

Representação do Conhecimento é o estudo do pensamento como um processo computacional, o uso de DL e seus mecanismos de inferências são propícios para este

trabalho, e deseja-se usá-los para representar e explicar bases de conhecimentos em domínios específicos. No desenvolvimento deste trabalho, são apresentados os estudos realizados na utilização de DL na formalização da semântica de textos em linguagem natural, para, posteriormente, empregar-se provadores de teoremas para raciocinar sobre a informação contida nos textos originais.

2. Fundamentação Teórica

Processamento de linguagem natural (PLN) é uma sub-área da inteligência artificial e da linguística que estuda os problemas da geração e compreensão automática de línguas humanas naturais, no qual através deste trabalho, estabelece-se uma relação com a lógica descritiva, para representação formal de conhecimento.

Sistemas de geração de linguagem natural convertem informação de base de dados de computadores em linguagem normalmente compreensível ao ser humano, de forma lógica e coerente, contudo, convertem ocorrências de linguagem humana em representações mais formais, mais facilmente manipuláveis por programas de computador, ou como mencionado no título deste trabalho, processador de linguagem natural. Processamento de linguagem natural é um método atrativo para interação homem-máquina, que pode ser aplicado a problemas de maiores níveis de ambiguidade e complexidade das inserções lógicas. Com base nesta afirmação, pode-se relatar que o uso de uma linguagem lógica para formalização de um texto normativo é indispensável e importante em todo o processo, bem como, no auxílio da construção da base de conhecimento.

Precisamos inicialmente estudar o que significa compreender, conceito-base de PLN. A compreensão é o ato de transformar uma forma de representação em outra, que seja significativa para o ambiente em questão e que permita o mapeamento para um conjunto de ações apropriadas, tanto para fins de armazenamento da informação, quanto para a tomada de algum tipo de decisão.

Para nosso trabalho, *Description Logic* - DL foi utilizada, tem sido amplamente utilizada em Bases de Conhecimento, emergindo como um conceito capaz de compreender e representar o conhecimento. DL não é um conceito novo, os primeiros estudos foram em meados da década de 70, cuja época, foi ponto marcante para estudiosos da área, pois surgiu o conceito de DL, que se define basicamente em um nome, que se refere a qualquer uma das várias linguagens lógicas, comumente utilizadas em Representação do Conhecimento [Davis *et al.*, 1993].

Portanto, DL são conjuntos de formalismos de representação do conhecimento de um domínio. Na Figura 1, é mostrado a estrutura de um sistema em DL, primeiramente, define-se os conceitos relevantes ao domínio, e utilizam estes conceitos para especificar as propriedades de objetos e indivíduos do domínio, criando a descrição do domínio. Uma lógica descritiva é formada por uma linguagem descritiva, que é utilizada para definir como os conceitos e como os papéis são formados.

3. Estado da Arte

Como mencionado anteriormente, processamento de linguagem natural é um método atrativo para interação homem-máquina, sistemas mais antigos como SHRDLU, trabalhava com *'blocks worlds"* restritos, com vocabulários restritos, levando pesquisadores a um excessivo otimismo, que mais tarde foi superado quando o sistema foi aplicado a problemas

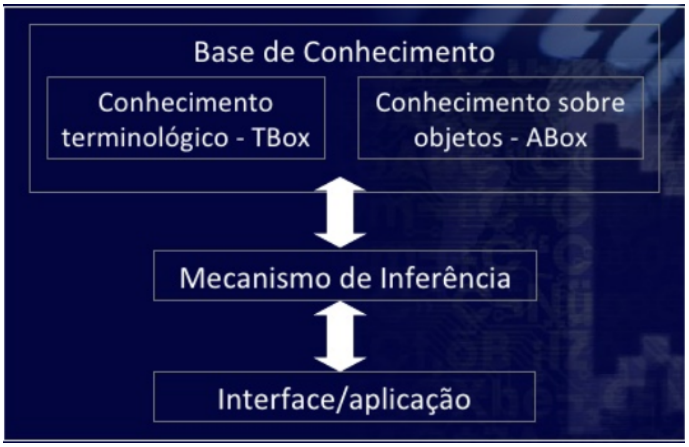


Figura 1. Estrutura de um sistema em DL

mais realistas, envolvendo ambiguidade e complexidade. No que se trata de processadores de linguagem natural, nosso trabalho se destaca e apresenta inovação, com utilização de DL e Ontologia OWL, que será detalhada nas páginas seguintes.

4. Desenvolvimento

O processador de linguagem natural consiste em representar o conhecimento de um texto em linguagem descritiva (DL) , assim estabelecendo um relacionamento com a semântica do domínio pré-definido. Dessa forma, a área de Processamento de Linguagem Natural (PLN) possui ligações significativas com o desenvolvimento deste trabalho, assim como a definição de ontologias, que são usadas para capturar conhecimento sobre um domínio de interesse, descrevendo conceitos e relações do domínio. Um dos mais importantes conceitos em ontologias, é a Ontologia OWL *(Web Ontology Language)*, este é um padrão definido mais recentemente e monitorado pela W3C *(World Wide Web Consortium)*.

Sabendo-se que, para muitas organizações a formalização de textos baseados em informação é muito importante, é comum encontrar documentos de texto, para serem avaliados por um agente, para processamento e tomada de decisões. Assim sendo, apresenta-se um exemplo de formalização de textos normativos no domínio da Segurança da Informação (SI) e de extração de conhecimento destes textos de linguagem natural. Para isso, primeiro definem-se os conceitos relevantes de um domínio/terminologia e então, usando estes conceitos, especificam-se as propriedades dos objetos e indivíduos deste domínio. Porém, é necessário que haja uma validação dos controles de segurança, assim, a terminologia relacionada a SI, de forma breve, é: controle de segurança, políticas de segurança e padrões de segurança, ambos são formalizados como conceitos na ontologia.

Neste sentido, queremos verificar se um controle de segurança é a implementação de uma ação específica do ponto de vista lógico. Para tanto, se o Controle01 possuir a seguinte descrição da política de segurança:

O tráfico de rede para a administração remota do servidor Netware deve ser criptografado usando SSL. E se a ação ‘Configurar todo o sistema para criptografar conexões usadas para o acesso remoto ao sistema”fizer parte de uma política de segurança de uma organização, nomeada como Ação02. Pode-se inferir que Controle01 de fato implementa

Ação02. Isto é feito provando que o conceito representando o controle é subsumido pelo conceito representando a ação, isto é, que a fórmula DL Controle 01 $\subseteq Ação02 é provável$.

Para formalizar as declarações ou as afirmações Controle01 e Ação02 na ontologia em formas lógicas, as seguintes considerações são feitas: frases na voz ativa ou passiva, no modo declarativo ou imperativo, e sentenças que contém os mesmos termos relacionados, devem se resumir a uma mesma forma lógica. Observe suas respectivas representações formais:

Controle01 $\equiv$
1. $\ni Verbo.(Criptografar \cap$
2. $\ni Tema.TráficoDeRede \cap$
3. $\ni Instrumento.SSL \cap$
4. $\ni Objetivo.(AdministraçãoRemota \cap$
5. $\ni Tema.ServidorNetware))$

$Ação02 \equiv$
1. $\ni Verbo.(Configurar \cap$
2. $\ni Tema.Sistema \cap$
3. $\ni Objetivo.(Criptografar \cap$
4. $\ni Tema.(ConexãoDeRede \cap$
5. $\ni éInstrumentoDe.(AcessoRemoto \cap$
6. $\ni Tema.Sistema))))$

No qual o Verbo é uma propriedade fornecida a fim de relacionar o conceito de ação com os conceitos verbais apropriados; Tema é uma propriedade específica para representar o tema do verbo; Objetivo e Instrumento são outras propriedades que representam papéis temáticos na ontologia; e éInstrumentoDe é uma propriedade inversa de Instrumento. Observe que a sentença de controle foi convertida para a voz ativa, e esta atitude é tida como um artefato desejável de formalização. O exemplo acima, de formalização, foi retirado de [Amaral *et al.*, 2006]. A prova, pode ser obtida em [Rademaker *et al.*, 2007], de que Controle01 $\subseteq Ação02$.

Para seguir os objetivos específicos deste projeto visando a obtenção de resultados palpáveis e bem concretizados teoricamente, a ferramenta Protege OWL foi utilizada para análise da estrutura de DL, no que se refere ao uso prático da mesma.

A ferramenta Protege OWL é dividida em 3 categorias principais: Indivíduos (*Individuals*), propriedades (*Properties*) e classes (*Classes*), no qual é possível definir os indivíduos do domínio, as propriedades às quais estes indivíduos pertencem e fazem relação com um ou mais indivíduos, e por fim é determinada as classes que envolvem o domínio, estabelecendo uma estrutura hierárquica do sistema e evitando inconsistência dos dados envolvidos. A Figura 2, tem-se um exemplo de hierarquia de um sistema de Pizza, demonstrando a usabilidade da ferramenta Protege e a contribuição para o desenvolvimento do protótipo de processador de linguagem natural.

Contudo, podemos afirmar que é baseado em um modelo lógico que torna possível definir os conceitos de forma como são escritos. Possibilitando definir conceitos mais complexos a partir de conceitos simples. Para validação e evitar inconsistência na estru-

Figura 2. Hierarquia de Classes

tura hierárquica da ontologia desenvolvida.

Ontologia OWL é classificada em 3 espécies:

1. OWL Lite (sintaticamente mais simples)
2. OWL DL (baseia-se em Lógica Descritiva, passível de raciocínio Lógico)
3. OWL Full (Destinada a situações com alta expressividade, não é possível efetuar inferências)

Com o intuito de desenvolver o protótipo de um processador de linguagem natural foi feito um estudo das diversas linguagens de programação existente, a priori a linguagem de programação Python foi objeto de estudo deste trabalho. Através do *framework* EasyGUI, disponibilizado gratuitamente na internet, é possível criar ambientes gráficos na linguagem Python, porém com o intuito de encontrar a linguagem de programação mais apropriada para desenvolver o protótipo do processador de linguagem natural, este *framework* foi apenas estudado e ponderado seu pontos positivos.

Como forma de ampliar os conhecimentos em linguagens de programação e diversas ferramentas de desenvolvimento gráfico, a ferramenta QTCreator foi estudada e colocada em prática durante o desenvolvimento do protótipo. Disponibilizada gratuitamente na internet, permite o desenvolvimento mais rápido do sistema quando comparado com o *framework* EasyGUI. Porém, apresenta maior incompatibilidade ao gerar o código fonte em Python da interface gráfica previamente desenvolvida, assim a ferramenta QTCreator foi apenas objeto de estudo/análise.

A linguagem JAVA, foi tida como apropriada para o desenvolvimento do protótipo, tal afirmação se baseia fortemente no conhecimento prévio adquirido durante o curso de graduação de Ciências da Computação na Universidade Federal de Goiás, a IDE (*Integrated Development Environment*) NetBeans na versão 7.2 foi um instrumento de trabalho para utilização linguagem JAVA. Na Figura 3, tem-se a representação da interface do Tradutor desenvolvido.

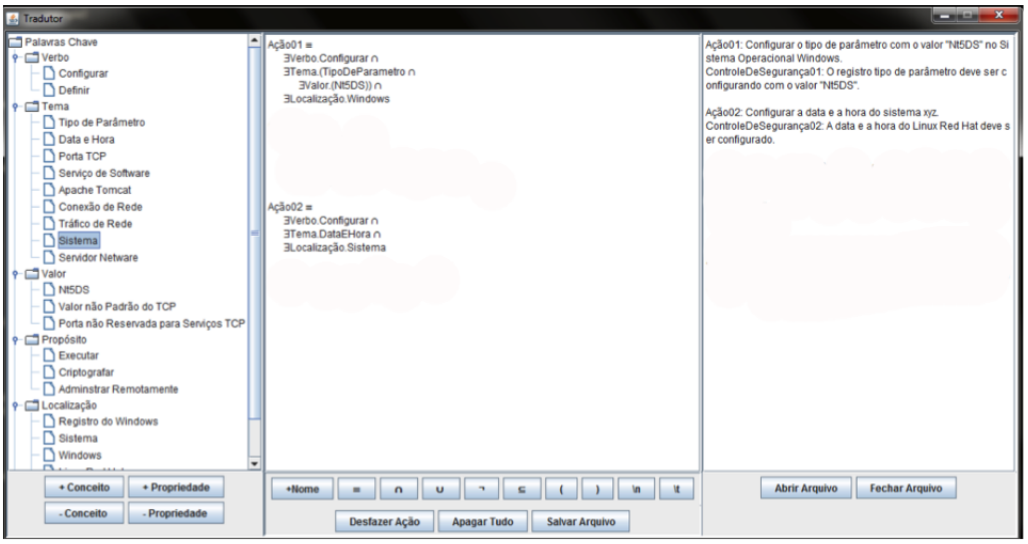


Figura 3. Protótipo do Processador em JAVA

4.1. Funcionalidades

O Processador de Linguagem Natural é dividido em 3 partes, a Figura 4 refere-se a Parte 1 do processador de linguagem natural, como mostrado abaixo:

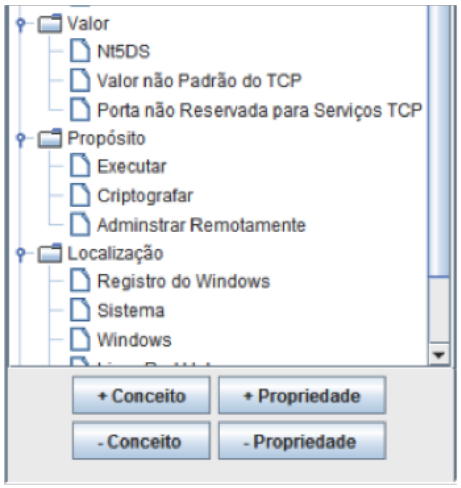


Figura 4. Parte 1 do Protótipo

Na Parte 1 do sistema, é possível adicionar com base no domínio de interesse, Conceitos e Propriedades. Temos: o Conceito Valor, para adicionar um Conceito basta clicar no botão +Conceito, e as propriedades que este conceito apresenta no domínio são, Nt5DS, Valor não Padrão de TCP, Porta não Reservada para Serviços TCP, para adicionar propriedades basta clicar no botão +Propriedade. A Parte 2 do sistema permite ao usuário inserir no processador de linguagem natural o problema descrito, salvo em arquivo texto. Para isto, basta clicar no botão Abrir Arquivo (Figura 4) e em seguida localizar no HD do computador o arquivo texto que contém a descrição do problema. A Parte 3, é destinada a formalização em DL, do problema inserido anteriormente, existe uma barra de símbolos lógicos na parte inferior da tela, tais como: +Nome (para adicionar nome ao problema), = (igualdade), $\cap(interseção), \cup(união), \neg(negação), \subset (estácontido)$.

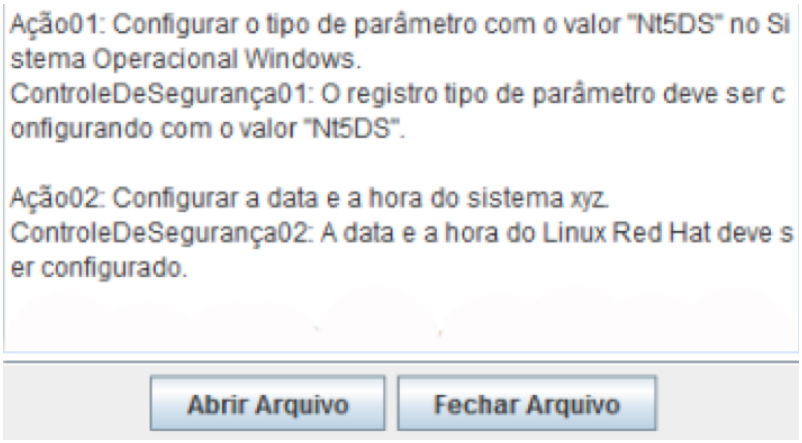


Figura 5. Parte 2 do Protótipo

Figura 6. Parte 3 do Protótipo

5. Resultados

As bases de desenvolvimento da área de Inteligência Artificial se encontram principalmente na noção teórica de "máquina de Turing' e na ideia de que "Pensar é computar', proposta pelo matemático, lógico e cientista da Computação Alan Turing. Os estudos de Turing contribuíram para o desenvolvimento da parte da Lógica relacionada com a análise simbólica do raciocínio [Tassinari, 2011].

Tendo-se um sistema de Base de Conhecimento, que armazena informações que descrevem o domínio através do formalismo de uma lógica, podem-se utilizar algoritmos de inferência que permitam a extração de conhecimentos implícitos e não percebidos pelo ser humano. Os raciocinadores lógicos podem incorporar mecanismos de inferência e utilizar diversos algoritmos, e são eles que garantem que o novo conceito incorporado à base de conhecimento, faça sentido e não cause nenhuma contradição nos conceitos já definidos no domínio [Mertins, 2011]. Dessa forma, processador de linguagem natural pode ser utilizado no processo de validação e garantir que elas não afetem de forma negativa, produzindo inconsistência e ou contradições na base de conhecimento.

Portanto, o uso de uma linguagem lógica para formalização de um texto normativo é indispensável, e nada mais justo que a utilização das lógicas de descrição que, em contraste com outros sistemas de representação, são equipadas com uma semântica formal e bem definida. Assim sendo, um Processador de Linguagem Natural se mostra, na prática, importante em todo esse processo e no auxílio da construção da base de conhecimento.

6. Conclusões

A partir da lógica descritiva, pode-se modelar o raciocínio humano, partindo-se de frases declarativas (ou proposições), caracterizando-se como o elo na automatização do processamento da informação. Na área de IA, por exemplo, a representação formal garante o correto raciocínio, expressando-se, assim, o alto poder de usabilidade de um raciocinador lógico. Logo, é de grande importância o estudo da lógica, pois o cérebro humano ter seu comportamento regido por relações lógicas. A lógica formal, mantém o rigor matemático necessário para modelar situações e analisá-las formalmente, desde então é aplicada e utilizada na validação sintática e semântica de softwares, na geração e otimização de códigos e na área de segurança em sistemas de informação.

Em suma, a lógica é uma área de grande êxito e se faz necessário a implementação de estratégias mais sofisticadas que envolvam raciocinadores, cujo objetivo é gerar formalização lógica em diferentes domínios de conhecimento. Há a interação com pesquisadores e alunos de pós-graduação do Laboratório de Técnicas e Métodos Formais (TecMF) do Departamento de Informática da PUC-Rio, que pesquisam tópicos relacionados a Representação do Conhecimento e Lógica de Descrição (DL), contando com a interação interinstitucional, que contribui para uma maior sinergia nos resultados.

Contudo, foi construído um Processador de Linguagem Natural, que mesmo classificado como protótipo, caracteriza-se como um refinamento do estudo sobre o tema, sendo possível obter resultados que, embora pareçam simples, são considerados um salto no estudo da automatização do conhecimento e do raciocínio lógico.

Referências

Davis, Randall. Shrobe, Howard. Szolovits Peter. (1993). *What is a Knowledge Representation?* AI Magazine, Volume 14, Number I.

Amaral, F. N., Bazílio, C., Silva, G. M. H., Rademaker, A., Haeusler, E. H. (2006). *Ontology-based Approach to the Formalization of Information Security Policies*. Em: 10th IEEE *International Enterprise Distributed Object Computing Conference Workshops* (EDOCW06).

Rademaker, A. R., Amaral, F. N., Haeusler, E. H. (2004). *A Sequent Calculus for ALC*. Monografias em Ciência da Computação. Em: 25/07. Pontifícia Universidade Católica do Rio de Janeiro. Setembro.

Tassinari, R. P., Gutierre, J. H. B. (2011). Lógica como Cálculo Raciocinador. São Paulo, 2011

Mertins, L. E. (2011). Extensão Virtual do Mundo Real: Integração Semântica e Inferência. Dissertação de Mestrado em Ciência da Computação. Universidade Católica de Pelotas. Pelotas.

Uso do NXTCAM V-4 Como Tecnologia Assistiva Para Deficientes Visuais

Luiz G. Dias[1], Luanna L. Lobato[1]

[1]Departamento de Ciência da Computação – Universidade Federal de Goiás (UFG)
Catalão – GO – Brasil

gusttavodiias@gmail.com, luannalopeslobato@gmail.com

Resumo. *É descrito neste artigo a importância das tecnologias assistivas na vida de usuários com deficiência. Deste modo é desenvolvido um modelo de cão guia robótico fazendo o uso do kit LEGO Mindstorms NXT 2.0 e uma câmera de modelo NXTCAM V-4. É realizado um estudo de caso em cenários diferentes que fizeram o uso de Wayfinding para comprovar a funcionalidade do robô.*

1. Introdução

Pessoas encontram diariamente obstáculos ou barreiras para ter acesso a informação, para se deslocar ou se comunicar, principalmente aquelas que possuem algum tipo de dificiência (DISCHINGER, 2012).

De modo a minimizar problemas como este, foi criado o termo acessibilidade, que permite não só que pessoas com deficiências sejam inclusas na sociedade, mas também pessoas com habilidades reduzidas ou deficiências temporárias consigam acesso espacial e digital a locais e informações. Assim diversas tecnologias vem sendo desenvolvidas e aplicadas para promover a inclusão de pessoas com deficiência, que são denominadas tecnologias assistivas.

Visto que existem diferentes tipos de deficiência, podendo ser classificadas entre físicas e cognitivas, este trabalho é focado no desenvolvimento de uma solução tecnológica assistiva para deficientes visuais através da robótica assistiva. Para tanto foi desenvolvido um modelo de cão guia robótico voltado a ambientes internos padronizados espacialmente com a finalidade de guiar o usuário. Sendo assim, é apresentada da Seção 2 a Fundamentação Teórica, na Seção 3 é mostrado o Desenvolvimento do Trabalho, na Seção 4 é apresentado o Estudo de Caso, e por fim na seção 5 são mostradas as Conclusões do trabalho.

2. Fundamentação Teórica

2.1. Acessibilidade

The subsection titles must be in boldface, 12pt, flush left.

A acessibilidade voltada a ambientes internos ou externos diz respeito a garantia de acesso, funcionalidade e mobilidade a todos os tipos de pessoa, independente de suas capacidades físicas e sensoriais. No que diz respeito a Acessibilidade digital, Castro(2009) mostra que tangente a acessibilidade voltada a ambientes físicos, a

Acessibilidade voltada a inclusão digital diz respeito ao acesso às plataformas tecnológicas visando a cidadania, pois o acesso de forma fácil e compreensível para diferentes grupos é essencial.

Deste modo pode-se definir acessibilidade como garantia de acesso a recursos e informação, independente da condição física ou cognitiva de quem a busca.

De modo a promover a acessibilidade, diversos tipos de técnicas e recursos são desenvolvidos, visando a inclusão, dentre as quais podem-se citar: próteses, auxiliadores de mobilidade e ambientes *Wayfinding*.

2.2. Tecnologias Assistivas

O termo tecnologia identifica todo o conjunto de recursos e serviços que contribuem para proporcionar ou ampliar habilidades funcionais de pessoas com deficiência (BERSH, 2008).

Cozinhe e Hussey (1995) definem Tecnologias Assistivas como uma ampla gama de equipamentos serviços, estratégias e práticas concebidas e aplicadas para diminuir os problemas encontrados pelos indivíduos com deficiências. Portanto todo e qualquer dispositivo ou técnica, destinada a assistir o deficiente em suas tarefas, se enquadra no domínio das Tecnologias Assistivas.

Sartoretto e Bersh (2013) definem recurso assistivo como todo e qualquer item fabricado em série ou sob medida, utilizado para aumentar, manter ou melhorar as capacidades funcionais das pessoas com deficiência. Os serviços são definidos como aqueles que auxiliam diretamente uma pessoa com deficiência em atividades como selecionar, comprar ou usar recursos, em contrapartida os serviços são aqueles restados profissionalmente à pessoa com deficiência visando selecionar, obter ou usar um instrumento de tecnologia assistiva.

2.3. Robótica Assistiva

Segundo Groothuis, Stramigioli e Carloni (2013), a Robótica Assistiva é um campo que cresce cada vez mais fazendo com que sistemas comerciais ou não sejam produzidos com a finalidade de assistir usuários portadores de deficiências permanentes ou temporárias.

A Robótica Assistiva engloba sistemas robóticos com assistividade, visando auxiliar pessoas incapacitadas permanente ou temporariamente, em atividades simples do dia a dia, como exemplos pode-se citar cadeiras de rodas motorizadas, robôs enfermeiros, andadores robóticos dentre outros.

2.4. Wayfinding

Heuten et al. (2008) definem *Wayfinding* como uma forma de resolver problemas espaciais de orientação. Desta forma o termo *Wayfinding* sugere sistemas de orientação espacial, permitindo melhor interação entre o usuário e o ambiente.

Segundo Li e Klippel (2010), alguns aspectos devem ser considerados ao desenvolver ambientes e sistemas de navegação que utilizam tal técnica, dentre elas se destacam os diferentes graus de conhecimento do ambiente por parte do usuário e o *layout* do ambiente, pois cada pessoa possui um determinado grau de compreensão

sobre um ambiente, e a complexidade do *layout* varia de acordo com o espaço físico, como o tamanho por exemplo.

3. Desenvolvimento do Trabalho

Para a efetivação do desenvolvimento da Tecnologia Assistiva citada anteriormente, foi necessário a utilização do kit LEGO Mindstorms NXT 2.0, da câmera NXTCAM V-4 utilizada como visão, além da implementação. A seguir são descritos cada item utilizado para o desenvolvimento do protótipo

3.1. Kit LEGO Mindstorms NXT

Os kits LEGO Mindstorms NXT são kits de robôs programáveis lançados pela LEGO em julho de 2006 que possui software próprio (SALAZAR, 2008).

O kit utilizado é um kit LEGO Mindstorms NXT 2.0, amplamente utilizado em escolas e universidades pelo cunho pedagógico. É composto por 619 peças: 3 servomotores, quatro sensores, alguns cabos e peças de encaixar.

O kit acompanha também uma CPU denominada *brick*. O *brick* é responsável por processar os dados recebidos pelos sensores, e enviar ordens para os dispositivos de saída (como os motores).

3.2. A Câmera NXTCAM V-4

Para coletar dados do ambiente, foi utilizada a câmera NXTCAM V-4, onde a mesma foi utilizada para funcionar como visão do protótipo, e assim ser possível detectar determinados objetos sinalizadores no ambiente, para que o robô se locomovesse de um ponto inicial até o ponto destino. A Câmera é mostrada na Figura 1.

Figura 1, Câmera NXTCAM V-4 utilizada no trabalho.

A Câmera possui dois modos de rastreamento, sendo estes *Object Tracking* utilizado para rastrear objetos, e *Line Tracking* utilizado para seguir linhas demarcadas no ambiente, além de conseguir acompanhar até oito objetos diferentes e fornecer em tempo real as estatísticas de objetos e cor (LEGO, 2013).

3.3. A Implementação

Levando em consideração as características técnicas da Câmera utilizada, foi decidido utilizar o mapeamento do ambiente através do rastreamento de objetos e não rastreamento de linhas, pois este tipo de mapeamento pode proporcionar uma gama mais flexível de possibilidades se tratando de ambientes *Wayfinding*.

Segundo Dias et al. (2013), a linguagem de programação nativa do kit LEGO não é recomendada para implementações complexas, pois como a mesma é feita de forma gráfica utilizando blocos pré-programados, se forem utilizados muitos blocos, a legibilidade e organização do código acaba sendo comprometida.

Para resolver este problema, a programação foi organizada de forma análoga a orientação a objetos, onde foram criados além do arquivo principal, três blocos de extensão '.rbt'. Tais arquivos são denominados bibliotecas ou blocos personalizados, tendo em vista que trabalham dados advindos de outros blocos. É mostrada na Figura 2 a relação entre os arquivos criados para implementação.

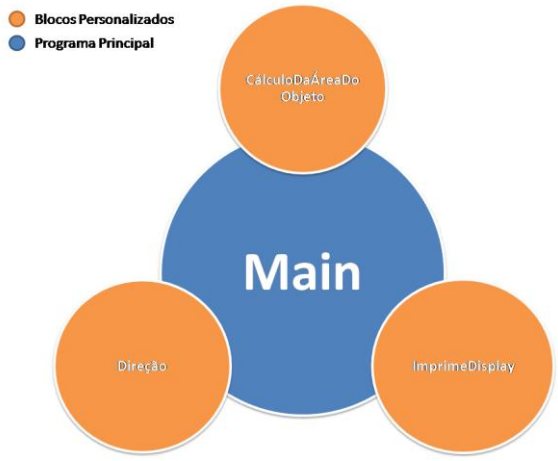


Figura 2, relação entre os arquivos criados para implementação.

Como mostrado na Figura 2, os blocos personalizados processam dados para que dados obtidos sejam obtidos e utilizados pelo bloco principal (na figura descrito como Main).

O bloco 'CalculoDaAreaDoObjeto' é responsável por calcular a área do objeto, onde a mesma é calculada de intervalos a intervalos, tendo em vista que a tendencia da área do objeto é aumentar de acordo com a proximidade do robô com relação ao objeto destino. O 'ImprimeDisplay' mostra no display do *Brick* a localização do objeto destino de acordo com as coordenadas X e Y, pois foi considerado no desenvolvimento do trabalho que o objeto destino pudesse se mover.

O bloco 'Direção' foi responsável por definir a direção da rotação dos motores, pois além do fator da escala do objeto, se o objeto se afastasse do protótipo, o bloco 'Direção' faria com que os motores se movessem para frente, visando aproximação, onde a faixa de parada foi definida como seis centímetros, ou seja, o protótipo se aproximava do objeto até que a distância fosse reduzida a seis centímetros. Por sua vez se o objeto se aproximasse do protótipo fazendo com que a distância definida como

faixa de parada diminuísse, o bloco 'Direção' fazia com que os motores se movessem para trás, evitando assim que ocorresse algum tipo de choque.

4. Estudo de caso

Para que a funcionalidade do modelo de Tecnologia Assistiva fosse testada, foi realizado um estudo de caso utilizando cenários diferentes. Tal estudo foi realizado com o auxílio do Departamento de Matemática Industrial, onde o autor atuou como participante do projeto de extensão denominado "Apoio a Capacitação no Uso das Tecnologias da Informação e Comunicação para a Juventude Rural – Uma proposta de inclusão digital para as comunidades Cisterna e São Domingos situados no município de Catalão". O estudo de caso foi realizado no Laboratório de Automação do DMI/UFG-CAC pelo fato do laboratório disponibilizar o material necessário, sendo estes o kit robótico utilizado para desenvolver a estrutura física do protótipo, e a câmera NXTCAM –V4, utilizada como visão do robô.

A tarefa realizada pelo protótipo constou em, de um ponto inicial o protótipo deveria localizar o objeto que demarca o destino final, se mover do ponto inicial indo de encontro ao destino final, parando a seis centímetros do objeto demarcador do destino final.

Os cenários no qual o protótipo foi testado foram:

- Cenário 1 – Curta Distância: utilizado para se obter os primeiros dados realizando o protótipo e o ponto de destino separados por curta distancia.
- Cenário 2 – Aumento da Distância: neste cenário a distancia entre o protótipo e o ponto de destino foi aumentada para que o alcance da câmera fosse testado.
- Cenário 3 – Iluminação Alterada: neste cenário o estudo de caso foi realizado com alteração no nível de luminosidade através de lâmpadas fluorescentes para que se fosse possível observar se fatores externos influenciariam no desempenho do robô.

A Figura 3 mostra as observações feitas durante o estudo no Cenário 1.

	Descrição
Local	Laboratório de Automação DMI/UFG-CaC
Material Utilizado	Protótipo montado através do kit robótico LEGO MINSTORMS NXT 2.0 utilizando a câmera NXTCAM-V4 como visão e uma esfera plástica da cor vermelha para demarcar o destino final.
Cenário	Inicialmente foi definido um intervalo de distância relativamente curto, de dois metros entre o protótipo e o objeto demarcando o destino. O destino foi demarcado com uma esfera de plástico de cor vermelha. O estudo de caso foi realizado com iluminação natural vespertina.
Variável Observada	Distância entre ponto inicial e destino final.
Resultado Preliminar	Nessas circunstâncias o protótipo localizou o objeto e foi ao seu encontro sem dificuldades, parando a seis centímetros de distância do mesmo, como previsto inicialmente. Se o objeto demarcando o destino se aproximasse, fazendo com que a distância entre o protótipo e o destino diminuísse, o mesmo se afastava a fim de evitar um choque.

Figura 3, observações no primeiro cenário.

Como visto na Figura 3, no primeiro cenário foram definidas as bases para a realização do estudo de caso onde o mesmo foi realizado definindo a distância como variável observada. Os resultados obtidos neste cenário foram compatíveis com os resultados esperados, o protótipo conseguiu localizar o objeto definido como destino final e ir ao seu encontro.

A seguir são mostradas na Figura 4 as observações realizadas do estudo no Cenário 2.

	Descrição
Local	Laboratório de Automação DMI/UFG-CaC
Material Utilizado	Protótipo montado através do kit robótico LEGO MINSTORMS NXT 2.0 utilizando a câmera NXTCAM-V4 como visão e uma esfera plástica da cor vermelha para demarcar o destino final.
Cenário	O segundo experimento constou em aumentar a distância entre o protótipo e o objeto que demarcava o destino final. A distância foi alterada para quatro metros entre o ponto de partida e o destino. O objeto foi demarcado com uma esfera plástica vermelha e a iluminação utilizada foi natural vespertina.
Variável Observada	Análogo ao Cenário 1, o Cenário 2 tem como variável principal a distância entre o ponto de partida e o destino, desta vez aumentando o valor da variável.
Resultado Preliminar	Como no Cenário 1, o protótipo também cumpriu com sua finalidade.

Figura 4, observações no segundo cenário.

Como notado na Figura 4, o cenário 2 foi realizado tendo como base o cenário 1, variando apenas na distancia entre o protótipo e o objeto definido como destino final. Como no primeiro cenário o protótipo realizou o trajeto sem dificuldades, localizando o objeto destino, uma vez que se moveu do ponto inicial até o ponto destino. De forma análoga ao Cenário 1, no Cenário 2 a variável observada foi a distancia.

Por fim, na Figura 5 são mostradas as observações obtidas do estudo no Cenário 3.

	Descrição
Local	Laboratório de Automação DMI/UFG-CaC
Material Utilizado	Protótipo montado através do kit robótico LEGO MINSTORMS NXT 2.0 utilizando a câmera NXTCAM-V4 como visão e uma esfera plástica da cor vermelha para demarcar o destino final.
Cenário	De forma análoga ao primeiro estudo, o terceiro cenário constou em realizar o estudo de caso com a distância entre o ponto de partida e o ponto destino em dois metros. O ponto destino foi demarcado com uma esfera vermelha, por sua vez a iluminação agora foi definida como artificial, fazendo o uso de lâmpadas fluorescentes além da iluminação natural vespertina.
Variável Observada	Diferente dos cenários anteriores, este cenário observa o funcionamento do protótipo tendo como variáveis principais a distância e a iluminação.
Resultado Preliminar	No Cenário 3 foi notada uma certa dificuldade para que o modelo de Tecnologia Assistiva localizasse o objeto demarcando o destino final, pelo fato de a esfera refletir a luz, entretanto o objeto foi localizado e o protótipo concluiu sua tarefa.

Figura 5, observações obtidas no Cenário 3.

De acordo com a Figura 5, é possível observar que diferente dos cenários 1 e 2, no Cenário 3 são observadas duas variáveis. Enquanto no Cenário 1 e Cenário 2 apenas a distância entre o protótipo e o objeto demarcando o destino final são observadas, no Cenário 3 são levadas em consideração a distancia e a iluminação artificial, possibilitada através de lâmpadas fluorescentes. Apesar do Cenário 3 contar com influência de fatores externos, e dificuldades de localização do objeto demarcador do destino final, o protótipo conseguiu realizar a tarefa proposta inicialmente.

5. Conclusão

Através da realização deste trabalho pôde-se perceber que diversas tecnologias podem ser utilizadas na construção de Tecnologias Assistivas. O propósito deste trabalho foi utilizar da robótica como meio provedor de acessibilidade para deficientes visuais, domínio que foi escolhido pela escassez de tecnologias que assistem o usuários em ambientes internos.

O estudo de caso comprovou a funcionalidade da Tecnologia Assistiva. No Cenário 1 e no Cenário 2, o estudo foi realizado sem influencia de fatores externos, favorecendo assim o *hardware* em questão. Já o Cenário 3 contou com a iluminação como fator externo para testar o desempenho do protótipo.

Tendo em vista os três cenários utilizados no estudo de caso pode-se concluir que a Tecnologia Assistiva desenvolvida cumpriu com a sua finalidade, que no caso foi guiar seu usuário em ambientes internos. Foi notado que questões relacionadas a fatores externos devem ser levadas em consideração durante a fase de projeto, pois os mesmos influenciam diretamente no resultado, como notado no estudo de caso do Cenário 3.

References

Boulic, R. and Renault, O. (1991) “3D Hierarchies for Animation”, In: New Trends in Animation and Visualization, Edited by Nadia Magnenat-Thalmann and Daniel Thalmann, John Wiley & Sons ltd., England.

Dyer, S., Martin, J. and Zulauf, J. (1995) “Motion Capture White Paper”, http://reality.sgi.com/employees/jam_sb/mocap/MoCapWP_v2.0.html, December.

Holton, M. and Alexander, S. (1995) “Soft Cellular Modeling: A Technique for the Simulation of Non-rigid Materials”, Computer Graphics: Developments in Virtual Environments, R. A. Earnshaw and J. A. Vince, England, Academic Press Ltd., p. 449-460.

Knuth, D. E. (1984), The TeXbook, Addison Wesley, 15th edition.

Smith, A. and Jones, B. (1999). On the complexity of computing. In *Advances in Computer Science*, pages 555–566. Publishing Press.

Model-Driven Development and Formal Methods: A Literature Review

Valdemar Vicente Graciano Neto[1,2]

[1]Instituto de Informática – Universidade Federal de Goiás (UFG)
Alameda Palmeiras, Quadra D, Câmpus Samambaia Caixa Postal 131 -
CEP 74001-970 - Goiânia - GO - Brazil

[2]Laboratório de Engenharia de Software (LabES)
Instituto de Ciências Matemáticas e Computação (ICMC)
Universidade de São Paulo - USP
Avenida Trabalhador São-carlense, 400 - Centro.
CEP: 13566-590 - São Carlos - SP - Brazil

valdemarneto@inf.ufg.br

Abstract. Background*: Model-Driven Development (MDD) has increased. However, MDD still has some lacks that becomes hard the entire adoption of the paradigm. On the other hand, Formal Methods (FM) has been applied in a lot of areas inside Software Engineering once they are mathematically based techniques for the specification, analysis and development of software systems, increasing software quality.*
Aim/Research Question*: This paper aims to answer this research question: How and how much MDD and FM have been associated in research? Answering that delivered a diagnostic report about this research field status.*
Method*: A literature review was conducted following Mapping Studies principles.*
Conclusions*: This research field has increased along the years, what shows promising results and cross-fertilizing benefits when these research fields are associated.*
Contribution*: An indication of research direction for Software Engineering area in the present and next years.*

1. Introduction

Model-Driven Development (MDD) has increased. It consists in a new prescriptive model of software development process [Pressman 2010, Graciano Neto and de Oliveira 2013] where models are the first class citizens [Sendall and Kozaczynski 2003, Amrani et al. 2012].

However, MDD still has some lacks that becomes hard the entire adoption of the paradigm to generate software in industry using exclusively this fashion, for instance:

1. the difficulties to automatically generate the entire software (a part is still manually generated),
2. the problems to validate model transformations once they are code as other programs and suffer from the same classic problems of Verification and Validation activities,

3. techniques and tools to check the models and assure their conformance against their respective metamodels, and so on.

On the other hand, Formal Methods has been applied in a lot of areas inside Software Engineering, specially in the critical systems area as avionics, health and militar, where errors can not be found.

Formal Methods are welcome once they are mathematically based techniques for the specification, analysis and development of software systems. Correctness of software systems can be improved by formalising different products and processes in the life-cycle, enabling rigorous analysis of the system properties [Oquendo 2006].

Thus this paper presents a literature review conducted following Mapping Study principles [Petersen et al. 2008, Kitchenham et al. 2010] to evaluate how Formal Methods could be associated to MDD. This could be a way to benefit and make it possible to construct and assure MDD artifacts quality and to become the use of Formal Methods more abstract, high-level, and human-readable.

The remainder of this paper presents some background in section 2, the research methodology in section 3, results in section 4, and some conclusions and future work in section 5.

2. Background

Several models are used to express the concepts of a knowledge domain for which a software is built. There are specific models for each phase of the software development process. The requirements model is used as input for discussion and production of architecture/design models. These models are considered for structuring the source code and the source code is the input for test cases specification. In fact, software development process is a succession of models transformations[Graciano Neto et al. 2010].

Model-Driven Development (MDD) is a software production process based on MDA (Model-Driven Architecture), a model-centered approach, where models, metamodels, and plugabble transformations are stored in *cartridges* that should be changed against a model transformer to produce many different kinds of software products in many technology flavors from a same source model.

MDE appeared after as an evolution of MDA and MDD to denote an complete process composed by methods, tools, technologies and principles that should be followed in the entire software development process to produce software as a sequence of model transformations.

MDE followed the natural evolution of any technology. It emerged as a coding activity in the software development life cycle. However, the principles started to be migrated by the software development team for the earliest moments of the software development life cycle as requirements engineering, software architecture definition and so on [Junior and Winck 2006], originating a totally new software development life cycle, as happened with aspects, agents, components, and tests, creating paradigms as Aspect Oriented Software Development/Engineering, Agent-Oriented Software Engineering and so on.

For simplicity, this paper uses MDD as the acronym to denote every sort of model-

driven technologies and acronyms.

MDD lives a deadlock. Tools and promising technologies as Eclipse Modeling Framework[1] and Kermeta [Manset et al. 2006] have been developed to support the MDD process. However, some studies have reported that MDD is not still mature [Westfechtel 2010] once some difficulties has been faced as:

1. The total automation is still a legend in most cases. Manual code addition is still necessary in a lot of model transformations;
2. Models are not enoughly expressive to cover every aspect of a software representation;
3. Model Transformations code is big and its production is still error-prone [Gonzalez and Cabot 2012]. Validation relies on the traditional software development techniques and the lack of validation techniques of models and model transformations are critical barriers to a wide industrial adoption [Baudry et al. 2010];
4. There is no guarantees that the software produced has quality and it is the expected result comparing it to one that could be manually produced.

The fact is that the representing software as models is an excelent idea, but there is no guarantees about its correctness, conformance, quality or rigor.

Formal Methods (FM) and *Model-Driven* do not often appear in the same sentence. Really, in a first moment we can see these words as representing disjoint areas.

However, if we reflect over this subject, we can envision that FM could strongly benefit MDD in some problems highlighted above as:

1. models could be **formally specified**, that is, models, metamodels and transformations could be specified using some formal method. This makes it possible to enforce and restrict the model expressiveness (through some formal lexical, syntatical and semantical value), becoming this totally measurable, allowing increasing or reducing the models according to the projects necessities;
2. models could be **formally verified**: again the models, metamodels and transformations specified in a formal way could be, either, formally verified. Additionally, this verification could be automatic once there are already tools and languages that do this kind of verification.

Additionally, the inverse situation can happen as well.

Formal methods allow a software engineer to create a specification that is more complete, consistent, and unambiguous than those produced using conventional or object-oriented methods. Set theory and logic notation are used to create a clear statement of facts (requirements). This mathematical specification can then be analyzed to prove correctness and consistency. Because the specification is created using mathematical notation, it is inherently less ambiguous that informal modes of representation [Pressman 2014].

FM have not been widely adopted in Software Engineering Industry (except Critical Systems, and Critical Embedded Systems) because [Pressman 2010]:

[1] http://www.eclipse.org/modeling/emf/

1. few software developers have the necessary background to apply formal methods, extensive training is required;
2. it is difficult to use the models as a communication mechanism for technically unsophisticated customers;
3. the development of formal models is currently quite time consuming and expensive.

But, considering MDD encompasses as least the high level and abstract software representation as models and software producing automation, MDD could help FM to increase level of abstraction in Formal Methods adoption, decreasing training time; to provide communication alternatives for customer and other stakeholders, and to automate the production of formal models synthesis from abstract formal models, minimizing time and cost.

Next section presents the research methodology we followed to investigate how these methods (FM and MDD) have been associated and reported in the software engineering literature.

3. Research Methodology

This paper aims to answer this research question: ***How and how much MDD and FM have been associated in research?*** Answering that delivered a diagnostic report about this research field status.

To investigate this research topic, a literature review was conducted following Mapping Studies principles [Petersen et al. 2008, Kitchenham et al. 2010]. This is an scientific investigation approach classified as an Exploratory Study once it has not the responsibility to compare any tools or methods, but just investigate and present a broad vision about some subject [Petersen et al. 2008].

A search was conducted using the string *Model-Driven Development and Formal Methods* in the ACM Digital Library[2] and Google Scholar[3]. Other scientific bases were not considered.

Two selection criteria were chosen: 1) Title Reading, and 2) Abstract Reading. Just papers and articles were considered in this research. After this step, the appropriate ones were included to a subsequent paper entire reading, and the inappropriate were discarded.

Once the search string was so broad, the amount of recovered results had been huge. Research validity discussions are done in a following section. Next section presents the results.

4. Results

Following the research methodology cited above, twenty-seven (27) papers were analysed. From those papers, eighteen (18) were considered appropriate and included while nine (9) were discarded because they were not treating the expected theme.

[2]http://dl.acm.org/

[3]http://scholar.google.com

Next subsections bring a quantitative analysis covering just numeric data, a qualitative analysis that discusses how MDD and MF have been associated in software engineering literature, and a discussion about the research validity.

4.1. Quantitative Analysis

Two aspects were observed during the data collecting: the kind of formal techniques focused on the papers, and the popularity of themes along the years.

Figure 1 presents the subject focused by the selected papers regarding to Formal Methods. Six focused just on Formal Verification. These papers did not focused their discussion on Formal Specification. Eight of them focused on Formal Specification and did not consider the formal validation of these formal models. Four of them reported both activities.

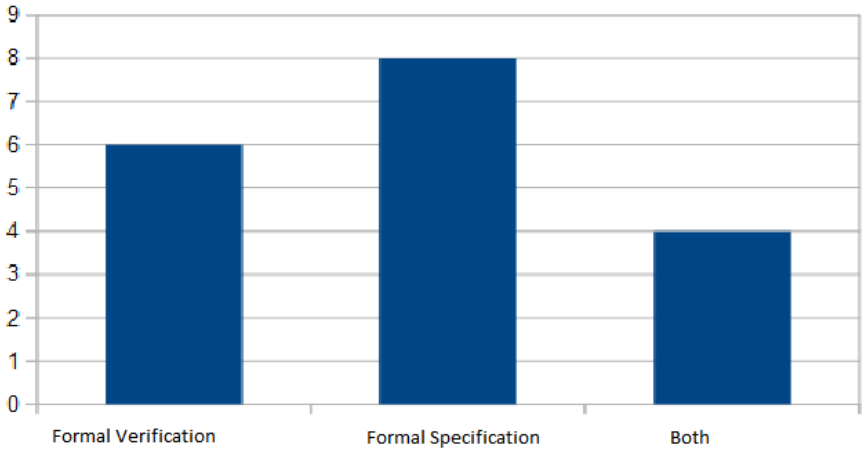


Figure 1. Research interest focused in papers.

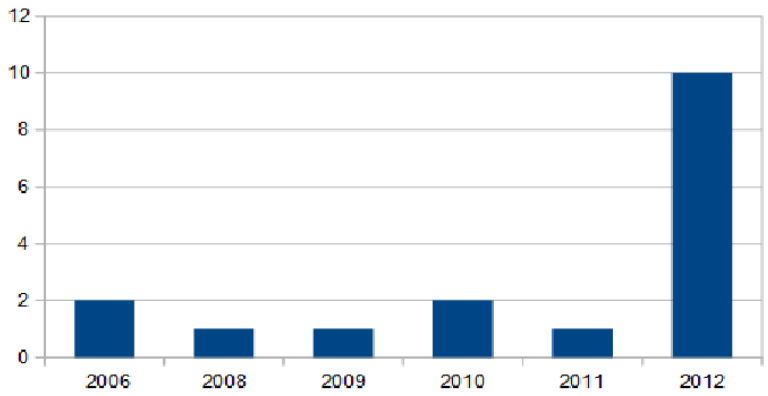


Figure 2. Popularity of the research topic in years.

Figure 2 analyses the popularity of this topic. It is evident that considering the amount of publications per year and the period covered by the selected results, the year 2012 has experienced a boom in this research topic, increasing in ten times the interest considerind with the preceding year (2011).

The only one paper written in 2013 recovered by the search was discarded. The research has been performed during 2013 November.

4.2. Qualitative Analysis

Between the works analysed, we found most of papers dealing with the association between FM and MDD as expected: some of them presented solutions to *formally specify* MDD artifacts, while others presented solutions to *formally verify* MDD artifacts.

Between the formalisms used we can highlight B, Z, Z3, formal diagrams as Statecharts, Petri Nets and UML extensions; EMF Modeling Operations (EMO), OCL, NuSMV, and so on.

About Model Transformations and FM, Three approaches have been used in the verification of Model Transformations [Lano et al. 2012] :

1. transformation code verifying (syntathic analysis);
2. mapping from transformation to a formalism in which the semantic analysis can be performed;
3. specifying transformations in a formalism for proof and implementation synthesis as a proof by construction.

Not so much papers from 2013 were recovered. One possible interpretation is that once the research has been performed in 2013, some conferences have been not indexed yet by the Scientific Search Engines while the research was being performed. However, the tendence of the results shows that 2013 could have even more papers dedicated to the highlighted research topic if the statistical increase keeps ascending.

No one result was found about how MDD could be used to increase formal methods level of abstraction to face those features that becomes the Formal Methods broad adoption in software engineering industry a hard task, except for those domains where it is imperative, as aeronautics, critical embedded systems, critical medical systems, military, and others.

This is a valuable contribution of this research, once it opens a new direction of research inside the conjunction of Formal Methods and MDD, introducing a new thread inside a research topic that can be considered, after these results, a hot topic and buzz words in Software Engineering research.

4.3. Research Credibility Discussion

This research was conducted under a strict available time, into the context of a PhD course in Formal Software Specification. This is the reason because not all the papers recovered were considered for the research.

Considering the nature of the study (exploratory) and the high-level and abstractedness of the search string, it is possible to suppose that the credibility of the study was not affected, once there was not an intention to be rigorous about the Scientific Search Bases amount, the papers number coverage or a strict comparing between mappings as it happens in Systematic Reviews.

The intention of this research is to give a panorama of how these research fields have been associated along the years, demonstrating a research tendence and an increasing interest.

The slice of papers cutted from the recovered results was specific. Considering the papers more aligned with the search string are recovered in the first positions of result in Information Retrieval Systems[Graciano Neto and Ambrosio 2009, Graciano Neto et al. 2009], the first ones were taken as more representative for the research question. So, with some percentual and statistic variation, we believe the first results can be considered the more relevant results for this research question.

5. Conclusions and Future Work

This paper presented results of an exploratory study conducted to answer the following research question: *How and how much Model-Driven Development (MDD) and Formal Methods (FM) have been associated in research*?

Answering that delivered a diagnostic report about this research field status, giving a panorama of how these research fields have been associated along the years, demonstrating a research tendence.

The main contribution of the paper is the opening of a new direction of research inside the conjunction of FM and MDD: the use of MDD to increase formal methods level of abstraction to face FM low adoption in Software Engineering industry. This can be glimpsed as a bew research direction for Software Engineering area in the present and next years.

It is possible to conclude that this research field has increased along the years, what shows promising results and cross-fertilizing benefits when these investigation areas are associated.

As future work, a more strict exploration could be performed over this theme, establishing a more restrict protocol and considering a larger slice of recovered papers, amplifying the research coverage. Furthermore, the new research direction identified can be explored, proposing new methods, models and techniques to deliver some building blocks to edify the basis for a more tangent, human, high-level and touchable formal methods process.

References

Amrani, M., Lucio, L., Selim, G., Combemale, B., Dingel, J., Vangheluwe, H., Traon, Y. L., and Cordy, J. R. (2012). A tridimensional approach for studying the formal verification of model transformations. *Software Testing, Verification, and Validation, 2008 International Conference on*, 0:921–928.

Baudry, B., Bazex, P., Dalbin, J.-C., Dhaussy, P., Dubois, H., Percebois, C., Poupart, E., and Sabatier, L. (2010). Trust in mde components: The domino experiment. In *Proceedings of the International Workshop on Security and Dependability for Resource Constrained Embedded Systems*, S&D4RCES '10, pages 2:1–2:7, New York, NY, USA. ACM.

Gonzalez, C. and Cabot, J. (2012). Atltest: A white-box test generation approach for atl transformations. In France, R., Kazmeier, J., Breu, R., and Atkinson, C., editors, *Model Driven Engineering Languages and Systems*, volume 7590 of *Lecture Notes in Computer Science*, pages 449–464. Springer Berlin Heidelberg.

Graciano Neto, V. V. and Ambrosio, A. P. L. (2009). A WordNet-based Search Engine for the Moodle Learning Environment. In *Proceedings of the International Conference EDUTEC*, EDUTEC '09, pages 1–10.

Graciano Neto, V. V., Ambrosio, A. P. L., and e Silva, L. O. (2009). Automatic Retrieval of Complementary Learning Material for Slide Presentations. In *Proceedings of the International Conference on Interactive Computer Aided Blended Learning*, ICBL '09, pages 1–8.

Graciano Neto, V. V., da Costa, S. L., and de Oliveira, J. L. (2010). Lessons Learned about Model-Driven Development after a Tool Refactoring (In portuguese). In *Proceedings of the VIII Annual Meeting of Computing*, ENACOMP '10, pages 68–75.

Graciano Neto, V. V. and de Oliveira, J. L. (2013). Evolution of an Application Framework Architecture for Information Systems with Model Driven Development(In Portuguese). In *Proceedings of the IX Brazilian Symposium on Information Systems*, SBSI'13, pages 1–12.

Junior, V. G. and Winck, D. V. (2006). *AspectJ - Aspect Oriented Programming with Java (In portuguese)*. Novatec, Sao Paulo, Brazil, 1 edition.

Kitchenham, B., Brereton, P., and Budgen, D. (2010). The educational value of mapping studies of software engineering literature. In *Proceedings of the 32nd ACM/IEEE International Conference on Software Engineering - Volume 1, ICSE 2010, Cape Town, South Africa,*, pages 589–598.

Lano, K., Kolahdouz-Rahimi, S., and Clark, T. (2012). Comparing verification techniques for model transformations. In *Proceedings of the Workshop on Model-Driven Engineering, Verification and Validation*, MoDeVVa '12, pages 23–28, New York, NY, USA. ACM.

Manset, D., Verjus, H., Mcclatchey, R., and Oquendo, F. (2006). A formal architecture-centric, model-driven approach for the automatic generation of grid applications. In *Proc. of the 8th Int. Conf. on Enterprise Inform. Systems*.

Oquendo, F. (2006). Pi-methodd: A model-driven formal method for architecture-centric software engineering. *SIGSOFT Softw. Eng. Notes*, 31(3):1–13.

Petersen, K., Feldt, R., Mujtaba, S., and Mattsson, M. (2008). Systematic mapping studies in software engineering. In *Proc. of ICEASE*, EASE'08, pages 68–77, Swinton, UK, UK. British Computer Society.

Pressman, R. (2010). *Software Engineering: A Practitioner's Approach*. McGraw-Hill, Inc., New York, NY, USA, 7 edition.

Pressman, R. S. (2014). Available in: http://www.rspa.com/spi/formal-methods.html. Access: February 2014.

Sendall, S. and Kozaczynski, W. (2003). Model transformation: The heart and soul of model-driven software development. *IEEE Software*, 20(5):42–45.

Westfechtel, B. (2010). A formal approach to three-way merging of emf models. In *Proceedings of the 1st International Workshop on Model Comparison in Practice*, IWMCP '10, pages 31–41, New York, NY, USA. ACM.

Simulação de Entrega de Fármacos Macro Encapsulados

Alessandro Machado Dahlke, Gustavo Stangherlin Cantarelli, Sylvio André Garcia Vieira

Sistemas de Informação – Centro Universitário Franciscano
97.010-032 – Santa Maria – RS – Brasil

alessandrodahlke@gmail.com, gus.cant@gmail.com, sylvio@unifra.br

Abstract. *The objective is the development of software that is capable of performing an approximation of the concentration of vitamin C in their bodies in free form as well as the analysis of other parameters by using a pharmacokinetic model which may reduce the number of animals involved in the testing lab. The model used is the monocompartiment, where it considers that the drug enters the body directly into the blood compartment. The use of the model allows comparing the results of administration of the same amount of drug intravenously and orally, and show that the concentrations achieved by the two methods are the same.*

Resumo. *O objetivo do estudo é o desenvolvimento de um software que seja capaz de estimar a concentração da vitamina c na sua forma livre em organismos, bem como a análise de outros parâmetros através da utilização de um modelo farmacocinético onde, poderá reduzir o número de animais envolvidos nos testes em laboratório. O modelo utilizado será o monocompartimental, onde o mesmo considera que a droga adentra ao organismo diretamente no compartimento sanguíneo. A utilização do modelo permitiu comparar os resultados da administração da mesma quantidade de droga por via intravenosa e via oral e mostrar que as concentrações atingidas pelos dois métodos não são as mesmas.*

1. Introdução

A farmacocinética é a área da farmacologia que realiza um estudo quantitativo do desenvolvimento temporal do movimento dos fármacos e seus metabólitos após sua administração em um organismo através da aplicação de modelos matemáticos [Goodman e Gilman 1996].

Os modelos matemáticos são utilizados através de modelos *in silico* (computacionais), estes que estão sendo integrados ao processo de planejamento e descoberta de fármacos a fim de se obter uma melhor predição do desempenho *in vivo* evitando desnecessários estudos de biodisponibilidade.

Normalmente, os programas são empregados na predição das propriedades físico-químicas de novos candidatos a fármacos, na simulação da absorção ao longo do trato gastrointestinal (TGI) e na abordagem do comportamento farmacocinético no organismo humano. Assim, este tipo de método pode ser útil nas predições da biodisponibilidade, em particular, de parâmetros como concentração plasmática máxima (Cmáx) e tempo em que

ocorre a concentração plasmática máxima (Tmáx) após administração [Terstappen e Reggiani 2001].

Considera-se a hipótese de que a concentração plasmática reflete a concentração da droga no local de sua ação. Para a predição da concentração da droga no organismo em relação ao tempo, será utilizado um modelo farmacocinético e dados empíricos para realizar simulações por via intravenosa e via oral.

Pretende-se desenvolver um software que seja capaz de estimar a concentração da vitamina c administrada por via intravenosa e via oral em organismos através da utilização de um modelo matemático. Com os resultados validados, pretende-se reduzir o número de animais utilizados em testes laboratoriais, minimizando os custos e permitindo a obtenção de uma maior gama de dados sem a necessidade de realizar novos experimentos.

Os resultados obtidos pelo software serão validados a partir de resultados empíricos presentes na literatura e a partir destes dados poderão ser criadas novas condições de simulação.

2. Modelos farmacocinéticos

A utilização de modelos matemáticos para simular o comportamento das drogas no organismo humano possibilita a obtenção de resultados sem a necessidade de experimentos com seres vivos. Este tipo de abordagem tem a vantagem de possibilitar a realização de simulações em condições críticas, que possam colocar em risco a saúde dos indivíduos, uma vez que experimentalmente seria inviável; e a vantagem de não necessitar de novos experimentos, baseando-se apenas na alteração de parâmetros para criar uma nova situação [Gallo-Neto 2012].

Com a utilização de um modelo numérico, por exemplo, basta inserir o valor desejado e o tipo de infusão que deverá ser feita. Se o mesmo procedimento fosse realizado em laboratório, além do risco gerado ao organismo em teste, seria necessária a utilização de certa quantidade de medicamento, gerando custo [Gallo-Neto 2012].

Os modelos farmacocinéticos, representados pelos modelos numéricos, necessitam de parâmetros que levam em consideração alguns fatores. Após a droga ser inserida no organismo, ela vai para o plasma sanguíneo e é distribuída ao longo do corpo. A velocidade com que o fármaco se propaga e é eliminado depende de como ele foi administrado no organismo, de como os tecidos o absorvem e de como é feita a sua eliminação do organismo [Gallo-Neto 2012].

Um dos principais objetivos dos modelos farmacocinéticos consiste em desenvolver um método quantitativo que descreva a concentração da droga no corpo como uma função do tempo. Apesar de serem artificiais e incompletos para representar um organismo, os modelos farmacocinéticos tem utilidade na interpretação dos processos de transporte e metabolismo das drogas. Os modelos mais comuns da farmacocinética utilizados para descrever a cronologia da variação das drogas são os modelos compartimentais [Silva 2006].

Existem modelos farmacocinéticos compartimentais, não-compartimentais e fisiológicos. Os modelos compartimentais são divididos em monocompartimentais (considerando apenas o plasma sanguíneo), bicompartimentais (incluindo além do plasma um compartimento periférico) e multicompartimentais (com a divisão dos órgãos e tecidos). Os modelos não-compartimentais descrevem a farmacocinética da droga

utilizando parâmetros do tempo e da concentração e os modelos fisiológicos descrevem a farmacocinética da droga em termos de parâmetros fisiológicos realísticos, tais como fluxo sanguíneo e coeficientes de partição tissulares. Os modelos compartimentais e não-compartimentais são modelos matemáticos abstratos, enquanto o modelo fisiológico se aplica a pacientes e situações clínicas reais [Silva 2006].

Neste trabalho o modelo monocompartimental será utilizado para determinar o tempo decorrido entre entrada de um medicamento no organismo e a situação de concentração na corrente sanguínea.

3. Modelo Monocompartimental

O modelo de um compartimento ou monocompartimental é o mais frequentemente utilizado na prática clínica dado a sua simplicidade matemática aliada a uma boa capacidade preditiva. É considerado o mais simples, inclui somente um compartimento e representa fármacos que após a administração, se distribuem para todos os tecidos atingindo rapidamente o equilíbrio em todo o organismo. Esta aproximação torna-se vantajosa na medida em que modelos mais complexos são de difícil aplicação devido à escassez de informação existente quando comparada com o número de parâmetros a determinar para a sua resolução [Bourne 2000].

Em um modelo monocompartimental, a administração pode ser por via intravascular, ou seja, toda a dose do fármaco entra imediatamente na circulação sistêmica e não ocorre processo de absorção (via intravenosa), ou por via extravascular, quando a etapa de absorção deve ser considerada (via oral) [Welling 1997] [Shargel 1999] [Dipiro 2002].

A concentração da droga é o critério principal para caracterizar e classificar os processos responsáveis pelo seu movimento de um local para outro no organismo. Se o processo farmacocinético de uma substância é descrito como sendo de primeira ordem ou exponencial, significa que a taxa de transferência ou de metabolismo da substância é proporcional à sua quantidade (concentração). Quanto maior a concentração, maior quantidade de droga sai em determinado tempo e quanto menor concentração, menor quantidade de droga é eliminada. Uma droga, por exemplo, que seja injetada por meio intravenoso, em dose única, desaparece do sangue, distribuindo-se e eliminando-se de acordo com a cinética de primeira ordem. Os exemplos de alguns processos fisiológicos que seguem a cinética de primeira ordem são: absorção, distribuição, *clearance* renal, e também, metabolismo das drogas [Silva 2006].

Na cinética de ordem zero, a taxa de metabolismo ou transferência de uma substância é constante e não depende da concentração da substância. Quando uma droga é administrada através de infusão intravenosa contínua, temos um exemplo de cinética de ordem zero [Silva 2006].

A Figura 1 exemplifica um modelo monocompartimental de administração por via oral, onde a droga é absorvida no trato gastrointestinal, é distribuída no corpo como sendo o único compartimento e é eliminada do organismo.

Figura 1 - Modelo de um compartimento oral. Adaptado de: [Bourne 2000].

Na administração de medicamentos por via intravenosa, o principio ativo é colocado diretamente na corrente sanguínea através de uma infusão por meio intravenoso, deixando-o livre do metabolismo de primeira passagem pelo fígado e demais degradações pelo trato gastrointestinal que poderiam ocorrer em uma administração por via oral [Bourne 2000].

A Figura 2 exemplifica o modelo de administração por via intravenosa, onde a droga é livre do processo de absorção e fica totalmente disponível na corrente sanguínea para distribuir-se e ser eliminada conforme a sua ordem.

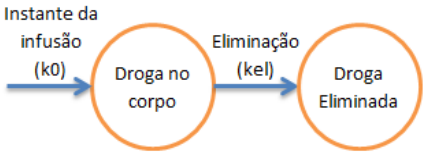


Figura 2 - Modelo de infusão intravenosa. Adaptado de: [Bourne 2000].

4. Projeto e Implementação do Software

Para implementação do software, foi realizada uma análise dos dados referentes ao fármaco, ao indivíduo em que será administrado o fármaco e dados utilizados pelo modelo matemático. Estes dados, bem como outros dados referentes a simulação são armazenados em um banco de dados para posteriores consultas e análise dos mesmos.

O software é configurado com parâmetros referentes ao indivíduo, ao fármaco e ao modelo utilizado. Permite que os resultados sejam apresentados de forma gráfica, podendo ser exportados em formato XML ou até mesmo salvos em uma base de dados. A Figura 3 mostra o modelo Entidade-Relacionamento do banco de dados utilizado pelo software, onde há uma tabela Farmaco para armazenamento de dados específicos de cada fármaco utilizado nas simulações, uma tabela Individuo para armazenar os dados de cada indivíduo a qual foi submetido a simulação, uma tabela Historico, para manter e disponibilizar os dados das simulações; e uma tabela Simulacao onde serão armazenados os resultados de cada simulação.

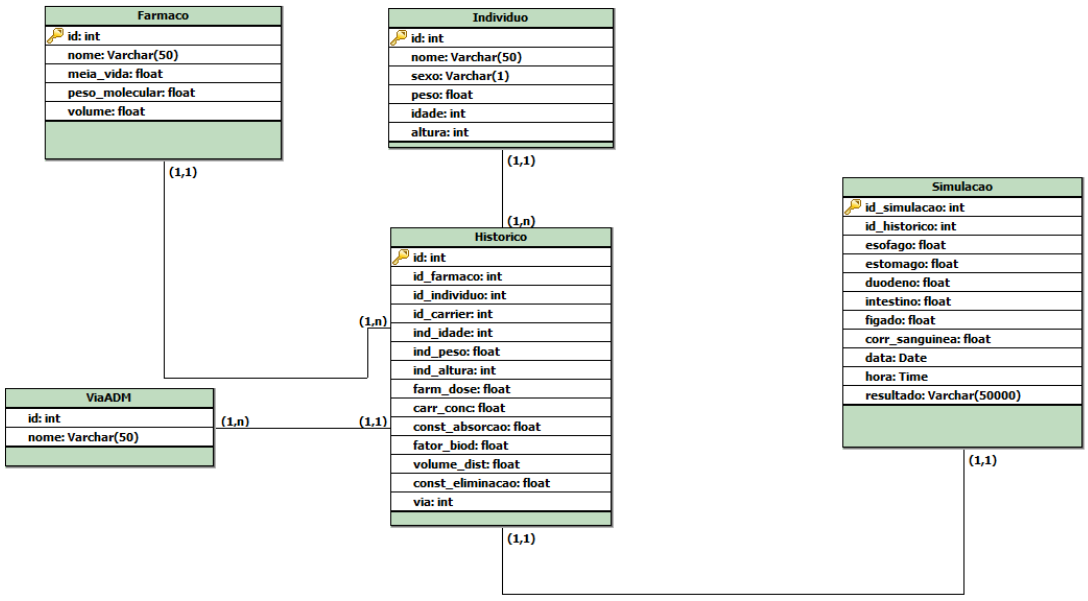


Figura 3 - Modelo entidade relacionamento de banco de dados

Neste trabalho serão abordados dois tipos de simulação, simulação por via intravenosa, onde o fármaco está livre do processo de absorção e simulação por via oral onde o fármaco deve ser absorvido antes de entrar na corrente sanguínea.

A infusão intravenosa tem como objetivo espalhar a droga no organismo de forma rápida almejando um efeito rápido garantindo que toda a dose entrará na circulação sanguínea, pois o fármaco não necessita ser absorvido pelo trato gastrointestinal e não está sujeito ao metabolismo de primeira passagem pelo fígado. Para prever as concentrações de fármaco em função do tempo em uma infusão intravenosa, é utilizada a cinética de primeira ordem, onde são consideradas as fases de distribuição e eliminação no organismo [Tozer 2009].

A Figura 4 mostra a equação utilizada para calcular as concentrações de fármaco (C) em cada instante de tempo (T).

$$Cp = Dose * e^{-kel * t}$$

Figura 4 - Concentração de droga após administração intravenosa [Bourne 2000].

A infusão oral é a mais comum e mais utilizada, a droga passa por várias etapas de desintegração do comprimido ou cápsula, dissolução da droga, difusão da droga através da membrana gastrointestinal, absorção ativa da droga através da membrana gastrointestinal até a sua entrada na corrente sanguíne. Os comprimidos e as cápsulas geralmente são formulados para liberar o fármaco logo após sua administração, visando apressar a absorção sistêmica. Essas são chamadas de formulações de liberação imediata. Outras formulações, as formas farmacêuticas de liberação modificada, foram desenvolvidas para liberar o fármaco a uma velocidade controlada, tendo como finalidade evitar o contato com o líquido gástrico (ambiente ácido) ou prolongar a entrada do fármaco na circulação sistêmica [Silva 2007].

A absorção oral de fármacos muitas vezes aproxima-se da cinética de primeira ordem, especialmente quando administrados em solução. Isso também ocorre com a absorção sistêmica de fármacos de muitos outros locais extravasculares, incluindo os tecidos subcutâneos e musculares. Nessas circunstâncias, a absorção é caracterizada por uma constante de velocidade de absorção, ka.

A concentração de fármaco na corrente sanguínea (Cp) após administração por via oral, acontece a partir da equação de primeira ordem, onde são considerados os parâmetros de biodisponibilidade (F), dose de fármaco administrado, constante de absorção (ka), volume de distribuição (V), constante de eliminação (kel) e tempo (t). A Figura 5 representa a equação utilizada para calcular a concentração de fármaco (Cp) em cada instante de tempo (t).

$$Cp = \frac{F \bullet Dose \bullet ka}{V \bullet (ka - kel)} \bullet \left[e^{-kel \bullet t} - e^{-ka \bullet t}\right]$$

Figura 5 - Concentração de droga após a administração oral [Bourne 2000].

A concentração plasmática máxima (Cmáx) representa a maior concentração sanguínea alcançada pelo fármaco após administração oral, sendo, por isso, diretamente proporcional à absorção. Desta forma, depende diretamente da extensão e velocidade de

absorção, e também da velocidade de eliminação, uma vez que esta se inicia assim que o fármaco é introduzido no organismo.

O tempo para alcançar a concentração plasmática máxima (Tmáx) tem íntima relação com a velocidade de absorção do fármaco e pode ser usado como simples medida desta. É alcançado quando a velocidade de entrada do fármaco na circulação é excedida pelas velocidades de eliminação e distribuição.

A área sob a curva de concentração plasmática *versus* tempo (ASC) representa a quantidade total de fármaco absorvido. É considerado o mais importante parâmetro na avaliação da biodisponibilidade, sendo expresso em quantidade/volume x tempo (mg/mL x h) e pode ser considerado representativo da quantidade total de fármaco absorvido após administração de uma só dose desta substância ativa. ASC é proporcional à quantidade de fármaco que entra na circulação sistêmica e independe da velocidade. Matematicamente, é obtida por cálculo através do método da regra trapezoidal.

5. Resultados obtidos

Devido ao pequeno número de trabalhos relacionados ao cálculo e a obtenção dos parâmetros da absorção de fármacos e a dificuldade de encontrar métodos matemáticos precisos que relacionem a fração da dose que é absorvida pelo organismo com a concentração plasmática, os resultados para as simulações foram obtidos a patir de parâmetros originados da literatura.

Foram realizadas simulações com a vitamina c por via intravenosa e oral. Percebeu-se na simulação por via oral que a absorção causa atraso e diminuição no pico de concentração; e na administração intravenosa, como não há absorção, o resultado foi a eliminação de primeira ordem da droga no decorrer do tempo.

A Figura 6 mostra a tela do sistema e os parâmetros utilizados para a simulação, onde o Ka é a constante de absorção, o F é o fator de biodisponibilidade, Vd é o volume de distribuição e Kel é a constante de eliminação.

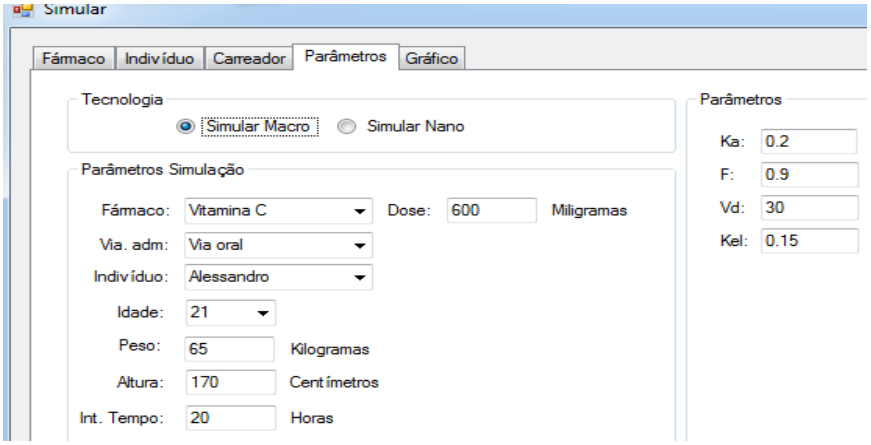


Figura 6 - Parâmetros farmacocinéticos utilizados na simulação oral

A Figura 7 mostra a concentração de fármaco em decorrer do tempo para as administrações intravenosa e oral da vitamina c. O fármaco entra no reservatório por um processo de primeira ordem e é eliminado do mesmo modo que o observado após dose intravenosa.

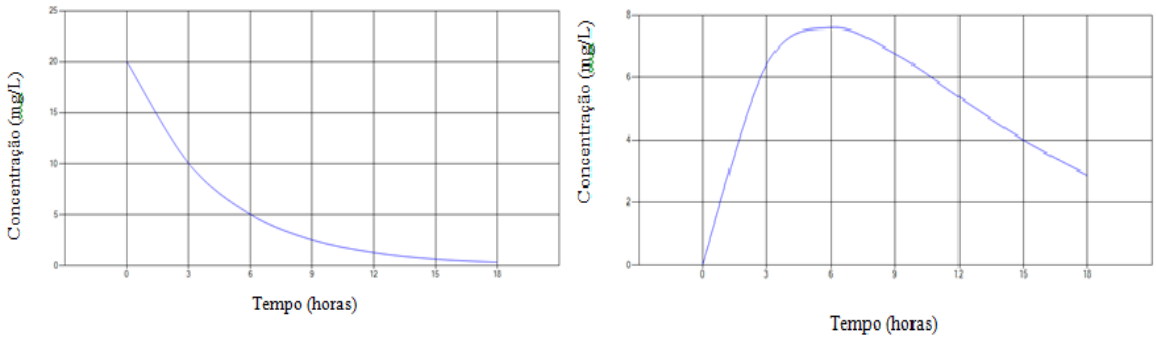


Figura 7 - Curva de concentração por via intravenosa e via oral

O pico de concentração plasmática após a administração oral é menor do que o valor inicial de concentração após a administração intravenosa de mesma dose. No primeiro caso, no tempo do pico certa quantidade de fármaco permanece ainda no local de administração, enquanto toda dose está no organismo imediatamente após a administração da dose intravenosa. Além do tempo de pico, a concentração plasmática excede aquela obtida após a administração intravenosa de mesma dose quando a absorção é completa.

Algumas dificuldades foram encontradas durante o desenvolvimento do trabalho, tais como a obtenção e entendimento das fórmulas matemáticas utilizadas para o cálculo das simulações. Há um grande estudo em relação à administração de medicamentos onde são avaliados os resultados referentes às concentrações na corrente sanguínea. Estes resultados, quando obtidos, tem sua origem a partir de dados referentes às coletas de amostras de sangue e de urina dos indivíduos. Através de métodos matemáticos é possível deduzir a quantidade de fármaco que tenha sido absorvida pelo organismo e estimar o quando foi eliminado, não sendo possível afirmar onde os processos de degradação, liberação e dissolução do fármaco ocorreram a partir de um comprimido ou cápsula utilizado na administração.

6. Conclusão

A possibilidade de redução do número de animais envolvidos é fundamental, assim como a possibilidade de visualização das taxas de absorções do fármaco em função do tempo tendo em vista as variáveis que agem sobre ela, como peso, altura e idade do ser que está recebendo o fármaco. Além disto, aspectos éticos e financeiros estão envolvidos e a demanda para realização de testes em laboratório está cada vez maior, dificultando a obtenção e estudos de biodisponibilidade.

O software foi desenvolvido para auxiliar no cálculo das concentrações e exibir a curva de concentração plasmática em função do tempo, dando-se os valores dos parâmetros ligados ao indivíduo e ao fármaco, como o volume de distribuição, a meia-vida de eliminação, a taxa de absorção e a fração absorvida. Com isso, o software em questão pode vir a ser uma ferramenta para se estudar a influência das variáveis farmacocinéticas no curso temporal dos níveis plasmáticos (Cp) de fármacos administrados *in vivo*.

7. Referências

Ansel, H. C. and Stoklosa, M. J. (2008) “Cálculos Farmacêuticos”. 12. Ed. Porto Alegre: Artmed. 451 P.

Bourne, D. W. A. (1994), "Mathematical Modeling of Pharmaceutical Data in Encyclopedia of Pharmaceutical Technology" Volume 9, Ed. Swarbrick, J. And Boylan, J.C., Dekker, New York, Ny.

Brunton, L. L., Lazo, J. S., Parker, K. L. Goodman and Gilman (2006), "As Bases Farmacológicas Da Terapêutica" (11ª Ed.). Rio De Janeiro: Mc Graw Hill.

Gallo-Neto, M., (2012), "Modelagem Farmacocinética E Análise De Sistemas Lineares Para A Predição Da Concentração De Medicamentos No Corpo Humano". São Paulo.

Goodman and Gilman, (2012), "As Bases Farmacológicas Da Terapêutica". 12. Ed. Porto Alegre: Editora McGraw Hill.

Gomes, M. and Reis, "A. Ciências Farmacêuticas: Uma Abordagem Hospitalar". São Paulo, Atheneu, 2000.

John and Golan, D. E, (2013) (S.D.). "Farmacocinética", P. 18.

Katzung, B.G. (1995) "Farmacologia Básica & Clínica", 6 Ed. Rio De Janeiro: Guanabara Koogan.

Moda, T. L.. (2007) "Desenvolvimento De Modelos In Silico De Propriedades De Adme Para A Triagem De Novos Candidatos A Fármacos", P. 82.

Rang, H. P., Dale, M. M., Ritter, J. M., and Flower, R. J. (2007) "Farmacologia" (6ª Ed.). Rio De Janeiro: Elsevier Editora Ltda.

Romão, M. I. (2012) "Simulação De Um Modelo Farmacocinético Para A Cisplatina", P. 48.

Rose. (1988). "Transporte De Ácido Ascórbico E Outras Vitaminas Solúveis Em Água." Biochim. Biophys. Acta V.947, P.335-366, 1988. Silva, Penildon, (2006) Farmacologia. 7. Ed. Rio De Janeiro: Guanabara Koogan.

Terstappen, G. C. and Reggiani A. (2006) "In Silico Research In Drug Discovery. Pharmacological Sciences." V. 22, P. 23-26.

Tozer, T. N., R, M (2009) "Introdução À Farmacocinética e À Farmacodinâmica - As Bases Quantitativas da Terapia Farmacológica".

Tubic, M.; Wagner, D.; Spahn-Langguth, H.; Bolger, M.B. and Langguth, P. (2006) "In Silico Modeling Of Non-Linear Drug Absorption For The P-Gp Substrate Talinolol And Of Consequences For The Resulting Pharmacodynamic Effect. Pharmaceutical Research." V. 23, P. 1712-1720.

Wagner and Nelson, E. (1964), J. "Pharm. Sci".

Em algum lugar além da diversão

Ivan da Silva Sendin[1]

[1]Faculdade de Computação
Universidade Federal de Uberlândia
sendin@facom.ufu.br

Abstract. *This work reviews the concept of Games With a Purpose (GWAP) and presents a library for a popular game: Plants Vs Zumbies. This library is intended to be used as teaching ressource in Artificial Inteligence courses.*

Resumo. *Este artigo apresenta uma revisão do atual estado das possibilidades oferecidas pelos jogos eletrônicos que vão além da simples diversão: jogos com propósito (GWAP, do inglês* Games With a Purpose*). Ainda apresentamos uma biblioteca para um popular jogo que pode ser usada como recurso didático nos cursos de algoritmos, otimização e inteligência artificial.*

1. Introdução

A indústria dos jogos eletrônicos é um indústria multibilionária [Rishe 2011]. Desde os seus primórdios ela anda de mãos dadas com o estado da arte do desenvolvimento tecnológico e científico, muitas vezes promovendo-os. Este artigo apresenta os recentes desenvolvimentos desta indústria que buscam algo além do simples entretenimento. Também apresentamos uma biblioteca para o desenvolvimento de um jogo similar ao *Plantas Vs Zombis*, da empresa PopCap[1]. A intenção desta biblioteca não é competir com o jogo original, seja em jobabilidade ou diversão, mas sim ser usada como ferramenta de ensino para os cursos de Ciência da Computação nas disciplinas como Algoritmos, Estrutura de Dados, Otimização e Inteligência Artificial.

O artigo é organizado da seguinte maneira: na Seção 2 apresentamos os aspectos históricos que fazem um paralelo entre o lúdico e a ciência de um modo geral; também, apresentamos alguns exemplos de iniciativas atuais que unam a diversão com jogos eletrônicos e alguma aplicabilidade prática. Na Seção 3, a biblioteca para o jogo é apresentada e um exemplo de seu uso como ferramenta didática na disciplina de Inteligência Artificial. Na seção 4 apresentamos as nossas conclusões.

2. Os Jogos e a Ciência

A relação de simbiose entre o lúdico e a ciência vem de longa data. No Século XVII, Blaise Pascal, matemático homenageado com uma linguagem de programação, desenvolveu a teoria da probabilidade baseada nos jogos de azar. O número i, a raíz da unidade negativa, foi criada em um processo de *competição* entre alguns matemáticos italianos [Mlodinow 2009].

Em [von Ahn and Dabbish 2008], seus autores introduzem o termo GWAP (*Games With a Purpose*, Jogos com Propósito) para classificar os jogos que ao serem jogados solucionam algum problema *paralelo* ao mesmo, este propósito de jogo é chamado também de *Computação Baseada em Humanos.* A seguir apresentamos alguns exemplos desta filosofia de motivação para jogos.

[1]`www.popcap.com`

2.1. Folding de Proteínas

A determinação da estrutura de proteínas dado a sua composição química é um trabalho computacionalmente complexo [Anfinsen 1973], existindo diversas abordagens usando os mais variados recursos da inteligência computacional [Donald 2005]. Em [Cooper et al. 2010a, Cooper et al. 2010b] é apresentado um jogo que consiste em montar proteínas seguindo restrições similares às restrições físicas que ocorrem nas proteínas reais. A montagem é iterativa e visual e o jogo é auto explicativo e didático pois instrui o jogador sobre conceitos de geometria molecular. Usando esta abordagem, milhares de usuários dispondo de tempo para o jogo obtiveram resultados expressivos [Eiben et al. 2012, Khatib et al. 2011b, Khatib et al. 2011a].

2.2. Câncer e Naves Espaciais

Recentemente pesquisadores associados ao *Cancer Research* do Reino Unido, desenvolveram um projeto de detecção de anomalias genéticas associadas ao câncer de mâma [Childs 2014]. O projeto reduziu o problema da detecção de anomalias ao problema de pilotar uma nave espacial, assim os inúmeros jogadores do jogo ajudam a detectar tais anomalias. O jogo esta disponível para as plataformas Androide e Apple.

2.3. PacMan

Pesquisadores do departamento de Ciência da Computação da Universidade de Berkeley apresentaram em [DeNero and Klein 2010] um projeto para ser usado nos cursos introdutórios de Inteligência Artificial. O projeto é baseado no popular jogo PacMan e permite o uso de diversas técnicas de IA como classificação, busca multi-agente e filtros de partículas. Ao aplicar estas técnicas os alunos *ajudam* o personagem PacMan a obter êxito no labirinto. É possível fazer o *download* das bibliotecas relacionadas no site `http://inst.eecs.berkeley.edu/~cs188/pacman/`.

3. A Biblioteca PxZ

Uma biblioteca para o jogo PxZ foi projetada para possibilitar o desenvolvimento de jogos no estilo Plantas Vs Zumbis. Ela foi implementada usando a linguagem Python. Dentre as funcionalidades implementadas estão a criação de plantas, criação de zumbis e os controles sobre o cenário do jogo.

A parte gráfica e a interface para jogabilidade foi implementada usando a biblioteca *pyGame* [McGugan 2007, Shinners 2011], embora estas funcionalidades não sejam necessárias para a biblioteca cumprir o seu objetivo.

O cenário do jogo é um jardim bidimensional, o jogador escolhe uma posição para criar a planta, que é escolhida de um conjunto de plantas pré-determinadas. As plantas têm a função primordial de defesa, seja ativamente atirando balas contra os zumbis ou fazendo uma barreira que retarda o avanço dos mesmos. As plantas também são responsáveis pela produção de *sois*, que são a unidade monetária do jogo.

Os zumbis são os inimigos a serem combatidos. São criados no extremo direito do jardim em uma linha escolhida aleatoriamente. Eles caminham em direção a esquerda. Ao encontrarem com uma planta eles a consomem. Eles podem ser eliminados por balas que são disparadas por alguns tipos de plantas. O jogo acaba se algum zumbi chegar ao fim do jardim, no lado esquerdo da tela.

3.1. As Plantas

A decisão da criação das plantas fica sob a responsabilidade do jogar, seja ele um jogador humano ou um sistema de inteligência artificial. A criação de uma planta é condicionada à existência de recursos.

A classe base das plantas é a classe **Planta** e possui as seguintes propriedades:

Massa A resistência da planta esta associada a sua massa, isto é, a quantidade de mordidas que a planta suporta antes de sucumbir ao ataque;

Frequência Algumas plantas executam ações[2] com uma determinada frequência.

As seguintes implementações de plantas foram feitas:

Girassol Este tipo de planta é responsável pela produção de sois, que é a unidade monetário do jogo e permite a aquisição de novas plantas. A produção de sol é determinada pela propriedade frequencia.

Ervilha É uma planta de defesa, que atira ervilhas que eventualmente matam os zumbis. Os tiros são retos e percorrem apenas um linha. Sua frequencia é determinada pela propriedade Frequencia.

Batata É uma planta passiva, sem ação. No jogo original sua massa é grande e tem a função de barreira.

3.2. Os Zumbis

Os *Zumbis* são criados de forma independente das escolhas feitas pelo jogador com uma frequencia pré-determinada e a escolha da linha aleatória.

3.3. O Jardim

O cenário do jogo é definido na classe **Jardim**, que controla o estoque de Sois, a criação de novas plantas e as iterações entre os personagens: plantas, zumbis e balas.

4. Experimentos Computacionais

Para mostrar a viabilidade do projeto desenvolvemos um algoritmo genético para criar uma estratégia de ação para o jogo. Esta estratégia é simples e apenas determina a planta que deve ficar em cada uma das posições do jardim.

4.1. Modelagem

Como as plantas estão dispostas no *jardim*, que é implementado em uma matriz, cada posição desta matriz recebe uma planta. Cada coluna desta matriz é codificada como um cromossomo.

Um novo indivíduo é formado pelos cromossos do pai e da mãe - escolhidos de forma aleatória - e por mutação pontuais. Para cada geração, apenas 30% dos indivíduos são escolhidos para fazer parte do *pool* a ser sorteado para terem os seus cromossomos passados para a próxima geração.

[2]No jogo original, algumas plantas são reativas ao contexto do jogo, nesta implementação as plantas não são reativas.

Nome	Massa	Frequencia
Girassol	5	3
Ervilha1	10	2
Ervilha2	1	40
Batata	1	1

Tabela 1. Definições para as plantas usadas nos experimentos computacionais. A planta Ervilha2 possui uma massa menor que a Ervilha1, assim deve ser *comida* mais rapidamente pelos zumbis; e uma frequencia maior, disparando um menor número da balas por unidade de tempo. Espera-se que a Ervilha2 seja preterida em relação a Ervilha1 no decorrer do tempo.

4.2. Função Objetivo

Quando um zumbi chega ao limite esquerdo do jardim, o jogo encerra e é decretada a derrota do jogador. A função objetivo a ser maximizada nesta modelagem é o número de iterações que ocorreram até o fim do jogo.

4.3. Resultados

Para testar o modelo genético proposto criamos um conjunto de plantas descrito na Tabela 1. A planta *Ervilha2* é propositadamente inferior a sua semelhante *Ervilha 1*. Como no início do sistema temos um conjunto aleatório de plantas, espera-se um equilíbrio entre o número da cada uma das as plantas escolhidas. Com o decorrer das iterações a opção pela planta *Ervilha 2* deve diminuir e eventualmente desaparecer.

Em um teste com uma população inicial de 100 indivíduos, a evolução para 100 gerações é descrita na Figura 1. Na Figura 2 mostramos um indivíduo aleatório da população inicial e outro da população final. O indivíduo da população final apresenta mais *Girassol* e *Ervilha1*, como era de se esperar.

5. Conclusões e Trabalhos Futuros

Neste artigo mostramos que os jogos computacionais podem trazer benefícios que vão além do divertimento. Apresentamos uma biblioteca para o jogo do estilo Plantas Vs. Zumbis e uma implementação de um algoritmo genético que desenvolve uma estratégia de jogo. A implementação é propositadamente simplória e deve ser vista como um recurso didático para os alunos desenvolverem a sua própria versão e o professor pode promover competições entre as diversas implementações. Outra possibilidade é deixar a criação dos zumbis também como uma atividade a ser desenvolvida como trabalho nas disciplinas.

Referências

Anfinsen, C. B. (1973). Principles that Govern the Folding of Protein Chains. *Science*, 181(4096):223–230.

Childs, O. (2014). Download our revolutionary mobile game to help speed up cancer research. `http://scienceblog.cancerresearchuk.org/2014/02/04/`.

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., Leaver-Fay, A., Baker, D., Popović, Z., and Players, F. (2010a). Predicting protein structures with a multiplayer online game. *Nature*, 466(7307):756–760.

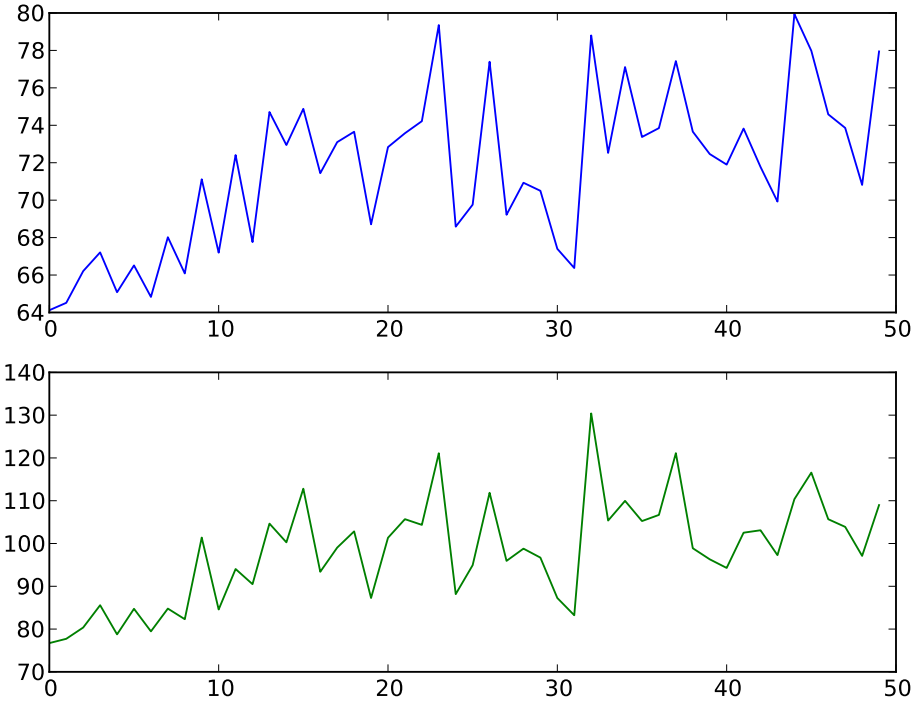


Figura 1. Evolução do número da média de iterações da populção.No gráfico superior, em azul, a média da população em geral e em verde a média da população escolhida para participar do processo de seleção genética.

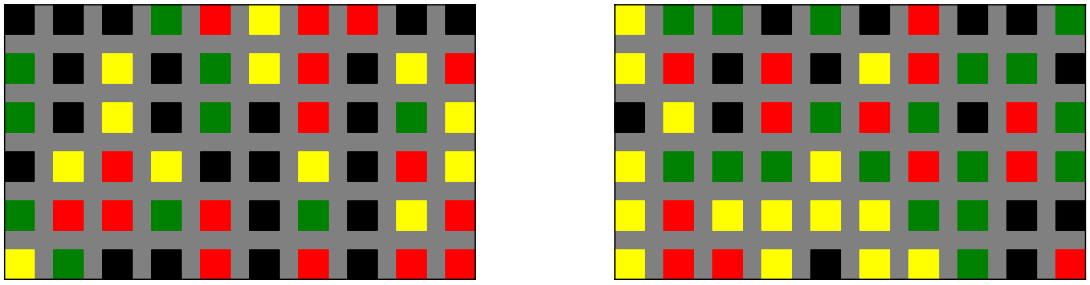
Cooper, S., Treuille, A., Barbero, J., Leaver-Fay, A., Tuite, K., Khatib, F., Snyder, A. C., Beenen, M., Salesin, D., Baker, D., and Popović, Z. (2010b). The challenge of designing scientific discovery games. In *Proceedings of the Fifth International Conference on the Foundations of Digital Games*, FDG '10, pages 40–47, New York, NY, USA. ACM.

DeNero, J. and Klein, D. (2010). The pac-man projects software package for introductory artificial intelligence. In *In proceedings of the Symposium on Educational Advances in Artificial Intelligence, Model Assignments Track*.

Donald, B. (2005). Algorithmic challenges in structural molecular biology and proteomics. In Erdmann, M., Overmars, M., Hsu, D., and der Stappen, F., editors, *Algorithmic Foundations of Robotics VI*, volume 17 of *Springer Tracts in Advanced Robotics*, pages 1–10. Springer Berlin Heidelberg.

Eiben, C. B., Siegel, J. B., Bale, J. B., Cooper, S., Khatib, F., Shen, B. W., Players, F., Stoddard, B. L., Popovic, Z., and Baker, D. (2012). Increased Diels-Alderase activity through backbone remodeling guided by Foldit players. *Nature biotechnology*, 30(2):190–192.

Khatib, F., Cooper, S., Tyka, M. D., Xu, K., Makedon, I., Popovic, Z., Baker, D., and Players, F. (2011a). Algorithm discovery by protein folding game players. *Proceedings of the National Academy of Sciences*, 108(47):18949–18953.



(a) Estado inicial do jardim, com as plantas escolhidas aleatoriamente.

(b) Estado final do jardim apresenta uma concentração maior de *Girassol* e de *Ervilha1*.

Figura 2. Posicionamento das plantas. Amarelo: Girassol, Verde: Ervilha1, Vermelho: Ervilha2 e Preto:Batata.

Khatib, F., DiMaio, F., Cooper, S., Kazmierczyk, M., Gilski, M., Krzywda, S., Zabranska, H., Pichova, I., Thompson, J., PopoviÄ, Z., Jaskolski, M., and Baker, D. (2011b). Crystal structure of a monomeric retroviral protease solved by protein folding game players. *Nat Struct Mol Biol*, 18(10):1175–1177.

McGugan, W. (2007). *Beginning Game Development with Python and Pygame*. Will McGugan, [New York].

Mlodinow, L. (2009). *The Drunkard's Walk: How Randomness Rules Our Lives*. Vintage Series. Pantheon Books.

Rishe, P. (2011). Trends in the multi-billion dollar video game industry. `http://www.forbes.com/sites/prishe/2011/12/23/trends-in-the-multi-billion-dollar-video-game-industry-qa-with-gaming-champ-fatal1ty/`.

Shinners, P. (2011). Pygame. `http://pygame.org/`.

von Ahn, L. and Dabbish, L. (2008). Designing games with a purpose. *Commun. ACM*, 51(8):58–67.

Análise no Uso do Controle de Concorrência no Acesso de Dados em MySQL e MariaDB

Hiago de Assis Silva[1], Flávio Ferreira Borges[2], Esdras Bispo Júnior[2], Paulo Afonso Parreira Júnior[2]

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

[2]Departamento de Ciências da Computação
Universidade Federal de Goiás – Campus Jataí – Jataí, GO - Brasil

hiagoassis@gmail.com, flavio@ufg.br, bispojr@ufg.br, pauloafpjunior@gmail.com

Abstract. *This paper presents a research on the performance in use of concurrency control on data access in MySQL and MariaDB, to analyze the behavior of these systems in terms of response time to control and manage this resource. And also show which is the performance gain the MariaDB in relation at MySQL. Thus it's possible to analyse the difference of the performance and verify which they servers testing obtains better gains with different loads.*

Resumo. *Este artigo apresenta um trabalho de investigação do desempenho no uso de controle de concorrência no acesso de dados em MySQL e MariaDB, visando analisar o comportamento destes sistemas em termos de tempo de resposta ao controlar e gerenciar esse recurso. E também mostra qual é o ganho de desempenho do MariaDB em relação ao MySQL. Dessa forma, é possível analisar a diferença de desempenho e verificar qual dos servidores testados obtém melhores resultados com diferentes cargas.*

1. Introdução

Com os avanços das organizações e o crescente aumento no volume de dados manipulados tornou-se necessário que os sistemas evoluíssem da mesma forma. Uma série de sistemas específicos foram desenvolvidos para suprir as diversas necessidades de um banco de dados, e essa tecnologia é nomeada Sistema Gerenciador de Banco de Dados (SGBD) [Elmasri *et al.* 2005].

Um SGDB é utilizado para manter controle sobre os dados e também sobre os programas que os acessam. Estes sistemas surgiram devido à necessidade de métodos de gerenciamento computadorizados mais eficazes. Dessa forma, permitem acesso aos dados armazenados em um banco de dados de maneira eficiente e precisa, atendendo as necessidades dos seus usuários [Silberchatz *et al*. 2006].

A comunicação entre usuários e SGBDs é possível graças a um padrão desenvolvido na década de 70, denominado *Structured Query Language* (SQL). A SQL é um padrão para SGBD relacionais e possibilita dentre outras funções, a requisição e manipulação aos dados por meio de consultas e atualizações, bem como a criação de estruturas para o armazenamento de dados (tabelas) [Elmasri *et al.* 2005], [Ramakrishnan *et al.* 2008].

Em geral, SGBDs permitem que várias requisições tenham acesso a um banco de dados ao mesmo tempo, ocasionando concorrência entre as transações[1]. Dessa forma, algum controle é necessário para garantir que transações concorrentes não possam prejudicar umas as outras. Portanto, existem diversos mecanismos de controle de concorrência que devem tratar algumas situações em que uma transação pode resultar em resposta errada, caso tenha sofrido interferência de outras requisições [Date 2003].

A técnica de bloqueio é o principal método utilizado para solucionar a concorrência entre transações [Silberchatz *et al*. 2006]. Normalmente, cada item de dado possui uma variável de bloqueio correspondente. E essas variáveis permitem obter informações sobre operações possíveis de ser aplicadas a estes itens de dados. Existem algumas formas diferentes de realizar bloqueio quanto ao controle de concorrência, entretanto, a aplicação dessas funcionalidades devem ser analisadas com cautela, pois podem implicar em problemas de desempenho, como exemplo, atraso excessivo para atender uma requisição [Elmasri *et al.* 2005].

Contudo, o objetivo deste trabalho é verificar o desempenho dos SGBDs MySQL e MariaDB, quando em uso de controle de concorrência no acesso aos dados, para configuração padrão e também por meio da implementação de bloqueio e desbloqueio (comando *GET_LOCK()).* Deseja-se verificar como se comportam, em termos de tempo de resposta o controle e gerência deste recurso nestes SGBDs.

A análise de desempenho permite obter métricas da eficiência de produtos, sobre alguma funcionalidade específica [Ciferri 1995]. Para avaliar o desempenho da concorrência nesse trabalho, é usada à técnica de *benchmark*. Segundo Vieira *et al.* (2005), o propósito do uso de *benchmarks* é apresentar formas padronizadas de avaliação do desempenho e confiabilidade de sistemas computadorizados.

Geralmente em *benchmark*, é definido primeiramente um conjunto de testes, os quais podem ser aplicados a diferentes sistemas, desde que apresentem funcionalidades semelhantes. Os resultados colhidos podem então ser analisados por uma metodologia de avaliação definida antecipadamente [Ciferri 1995]. O JMeter possui estas técnicas mencionadas e foi a ferramenta utilizada neste trabalho.

O Apache JMeter é uma ferramenta de código aberto e possui distribuição livre para aplicações de testes de sistemas, possibilitando sua aplicação também em rede distribuída. É baseada em *Java* e passível de extensão devido a *Application Programming Interface* (API) fornecida pela Apache. Após seu desenvolvimento, foi expandido para realizar testes em servidores *File Transfer Protocol* (FTP), servidores de banco de dados, e *JavaServlets* de objetos [Halili 2008].

Todavia, autores como Santos (2008) mostram a importância da realização de testes de desempenho e estresse para avaliar a qualidade de sistemas computacionais com a utilização da ferramenta de *benchmark* JMeter. Em Santos (2008) a avaliação é aplicada a SGBDs com intuito de verificar se existem possíveis perdas de desempenho e até mesmo a não disponibilidade de serviços.

A investigação sobre o desempenho entre MySQL e MariaDB se faz relevante pela preocupação em estar migrando aplicações que utilizavam o MySQL como seu

[1] Uma transação é uma unidade lógica de trabalho em banco de dados e envolve diversas operações.

gerenciador de dados para um novo SGBD, especificamente, o MariaDB. Este trabalho pode contribuir em pesquisas que buscam avaliar as características de um banco de dados e definir alterações para adequá-lo a um melhor padrão de funcionamento.

2. Trabalhos Relacionados

O objetivo que se tem com esses trabalhos é apresentar as discussões e problemas relacionados à SGBDs, e alcançar maneiras e métodos para realizar uma avaliação de desempenho entre *softwares*. De modo que seja fundamentada em um alto nível de imparcialidade para mostrar a relevância da realização desta análise.

Em Perseguine (2012), é apresentado o estudo do controle de concorrência com ênfase nas suas características em diferentes arquiteturas de banco de dados. Esse trabalho destaca a importância de manter o controle de concorrência em banco de dados e também como a manipulação de suas técnicas devem ser consideradas caso a caso. Mostra ainda, um estudo sobre o comportamento e evolução dos mecanismos de controle de concorrência, que buscam se adaptar aos pré-requisitos de novas arquiteturas.

O trabalho de Kyoung-Don *et al*. (2007) é um exemplo de como são manipulados sistemas computacionais que enfrentam sobrecarga de trabalho. Nesse trabalho, são abordados servidores de banco de dados sobrecarregados, em especial servidores de tempo real, ditos *real-time database* (RTDB). Para essas aplicações, o tempo de resposta ao (s) requisitante (s) é o ponto chave para seu sucesso.

O autor utiliza um tipo de tecnologia chamada *Chronos*, que pretende aliviar a sobrecarga sobre estes servidores. Esta ferramenta faz uso de algumas técnicas específicas, como exemplo, utiliza-se do bloqueio em duas fases para tentar manter o controle da concorrência sob as aplicações de tempo real utilizadas.

Em Pires *et al.* (2006) é realizado um comparativo de desempenho entre SGBDs de código aberto utilizando o *benchmark Open Source Database Benchmark* (OSDB), com objetivo de analisar as métricas geradas e sugerir possíveis melhorias em seus desempenhos. Este trabalho é útil para a presente avaliação por mostrar a importância de se realizar testes de desempenho em SGBD e como devem ser feitos por meio de ferramentas específicas para esse fim.

O trabalho de Santos *et al.* (2008) apresenta a importância da realização de testes de desempenho e estresse na qualidade dos produtos e também sobre o uso da ferramenta com técnicas de *benchmark* JMeter, de acordo com autor, essa ferramenta é apropriada para esse tipo de teste.

Em Silva (2013) é apresentado uma investigação sobre o desempenho do uso de visão em consultas ao banco de dados, permitindo identificar se o uso deste recurso ameniza a sobrecarga de requisições. Foram utilizados os SGBDs MySQL e PostgreSQL, e por meio dos resultados apresentados, foi possível notar que os dois gerenciadores obtiveram ganho de desempenho nas consultas quando em uso de visão, sendo este ganho foi superior no PostgreSQL. A ferramenta de *benchmark* utilizada por Silva (2013) foi o JMeter.

Por fim, a análise realizada por Schwenke (2012), apresenta os SGBDs MariaDB-5.5.24, MySQL-5.5.25 e Percona Server 5.5.24-26.0. Buscou-se neste trabalho, a realização de testes de desempenho por meio do *benchmark* Sysbench

OLTP. A partir de testes que envolvem operações de leitura e escrita no banco de dados, foram apresentados como resultados os valores de transações por segundo realizadas por estes SGBDs e também o tempo de resposta em que esses gerenciadores atenderam as requisições de diferentes quantidades de usuários enviadas pelo Sysbench.

Indicou-se nesta avaliação, com base nos resultados obtidos, que em relação a operações de leitura, não há muita diferença em termos de desempenho entre estes três SGBDs. Porém, em operações de escrita no banco de dados, afirmou-se que o MySQL e o MariaDB atenderam uma maior taxa de transferência de requisições por segundo.

Enfim, conforme apresentado, há uma necessidade de analisar e comparar tecnologias diferentes sob alguns aspectos específicos. Devem ser realizadas de forma automatizada e metodológica, de forma que certifique relevância e coerência na escolha dos *softwares* adequados para cada caso. Portanto, é evidente a importância da investigação proposta, pois a análise de desempenho retorna diretrizes que podem trazer ganhos a diversos sistemas.

3. Metodologia

A metodologia utilizada nessa análise é a *Goal Question Metric* (GQM). Ela indica que para realizar a coleta e avaliação de métricas é necessário primeiramente definir metas. Esse tipo de avaliação permite detalhar um sistema de mensuração, aplicado para um conjunto definido de características e um conjunto de regras que possibilitem interpretar os dados mensurados [Dal'osto 2004].

São detalhados abaixo, os pontos GQM dessa metodologia:

• Objetivo (*Goal*): O objetivo deste trabalho é verificar o desempenho do controle de concorrência no MySQL e MariaDB para os comandos SQL *select* e *update* variando o tamanho das consultas e atualizações e dessa forma influenciando diretamente no esforço do SGBD.

• Questões (*Question*): Qual é o desempenho em termos de tempo de resposta em que cada SGBD atende requisições de diferentes usuários simultaneamente? Para qual tipo de transação o SGDB tem um melhor desempenho?

• Métricas (*Metric*): Linha de 90% dos tempos de resposta fornecidas pelas JMeter e *Speedup*.

Linha de 90%: A métrica linha de 90% fornecida pelo JMeter, é uma medida adequada para testes de desempenho. Ela representa que 90% dos testes apresentaram valores até o tempo de resposta fornecido por essa métrica. Em avaliações, os valores obtidos sobre tempos de resposta sofrem grandes variações, ou seja, resultam em valores muito baixos e valores muito altos na mesma série de dados. Segundo Moffatt (2013) a média aritmética nestes casos pode não apresentar resultados eficientes.

Speedup: Ao considerar o impacto de alguma melhoria de desempenho em um sistema, ou de um sistema para outro, normalmente o efeito dessa melhoria é expresso em termos de aumento de velocidade. O *Speedup* (S) é uma expressão que mostra esse ganho de desempenho, considerando como a razão entre o tempo de execução sem a melhoria (T_{w0}) para o tempo de execução com a melhoria (T_w) [Murdocca 1999]. Para realizar este cálculo, considerou-se o MariaDB como sendo um sistema com melhorias, pelo fato de ser um SGBD recente e que oferece como alternativa ao MySQL, que

consequentemente foi o sistema a ser comparado. Dessa forma, o *Speedup* demonstra em determinadas situações qual é o ganho de desempenho do MariaDB em comparação ao MySQL.

$$S = \frac{T_{W0}}{T_W}$$

4.1. Descrição dos Testes

Os testes foram realizados em um servidor *DELL® PowerEdge* 2900, que utiliza o sistema operacional *CentOS* 6.4 final, com as seguintes configurações: 8 processadores de 2 núcleos *Intel® Xeon* E5410 2.33 GHz (cada núcleo), memória *RAM* de 8 GB, 4 HDs *SATA* 250GB 7.2k, localizado no laboratório de pesquisas em Ciências da Computação da Universidade Federal de Goiás, unidade de Jataí/GO.

A etapa inicial compreendeu na instalação dos bancos de dados MySQL 5.5.33 e MariaDB 5.5.34 no servidor. Para isso, foram utilizadas duas máquinas virtuais (*Oracle* VM *VirtualBox*), cada uma portando 3 GB de memória *RAM*, e utilizando 3 processadores (6 núcleos). O sistema operacional utilizado nas máquinas virtuais foi o *Debian* 7.2. *Wheezy.*

A escolha dos SGBD se deve a possibilidade de substituição do MySQL pelo MariaDB, em virtude da aquisição do MySQL por uma empresa que trabalha com softwares proprietários, além do fato de que grandes organizações estão migrando ou analisando migrar para esse SGBD de código livre. Já a escolha do sistema operacional *Debian*, foi devido ao maior reconhecimento da comunidade científica quanto à utilização de sistemas com derivação GNU/LINUX.

Para simular o mais próximo de um ambiente em que SGBDs são utilizados na prática, aplicou-se a ferramenta apache-jmeter-2.10 em uma máquina cliente conectada via rede ao servidor. Esse computador possui com as seguintes configurações: *Notebook Sony® Vaio*, processador *Intel®* i5, 4 GB de memória *RAM*, HD de 500 GB e utilizando sistema operacional *Debian 7.2. Wheezy*.

Para a realização dos testes, foi adotada uma base de dados de gerenciamento escolar, que contém 103 tabelas e 97 relacionamentos e com tamanho de aproximadamente 135MB, já igualmente predefinida para os dois SGBDs. Foram construídos seis comandos SQL para os dois servidores, sendo eles três *select* e três *update*, alguns destes possuem várias interações entre as entidades, de modo que exija elevado processamento de todo o sistema.

4.2. Carga de Trabalho

Buscou-se com as cargas de trabalho analisar o desempenho quanto à concorrência quando em uso de requisições de cargas que exigem pouco do sistema e também com cargas que necessitam de grande recurso computacional. Elas foram aplicadas aos dois SGBDs na seguinte ordem:

1. Foram executadas *select* de carga e *update* de carga baixa com 4, 8, 16, 32, 64, 128, 256, 512 e 1024 usuários simultâneos, no MariaDB e posteriormente no MySQL em configuração padrão.

2. Alterado o valor da variável *innodb_lock_wait_timeout*, foi realizado um *update* de carga alta no MariaDB e posteriormente no MySQL com 4, 8, 16, 32, 64, 128, 256, 512 e 1024 usuários simultâneos.

3. Foram executados *select* e *update* de carga alta com o uso de *GET_LOCK()* para o MariaDB e posteriormente no MySQL com 4, 8, 16, 32, 64, 128, 256, 512 e 1024 usuários simultâneos.

4. Por último, foram executados *select* e *update* de carga baixa, com o *pool-of-threads* ativo no MariaDB também com 4, 8, 16, 32, 64, 128, 256, 512 e 1024 usuários simultâneos.

4. Resultados Obtidos

Foram realizados diferentes testes, obtendo resultados da linha de 90% fornecida pelo JMeter. Com esses tempos de resposta, foram calculados a porcentagem da diferença entre os resultados para MySQL e MariaDB para cada tipo de teste. Também foram calculados os S*peedup* para verificar e se houve ganho de desempenho do MariaDB em relação ao MySQL para os mesmos casos.

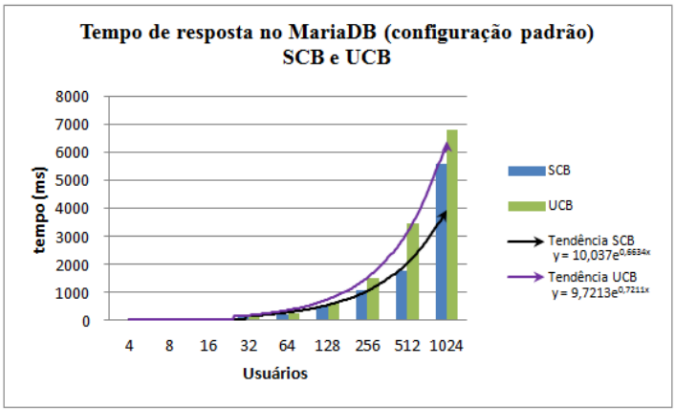


Figura 1. Tempo de resposta no MariaDB com a configuração padrão utilizando *select* com carga baixa e *update* com carga baixa.

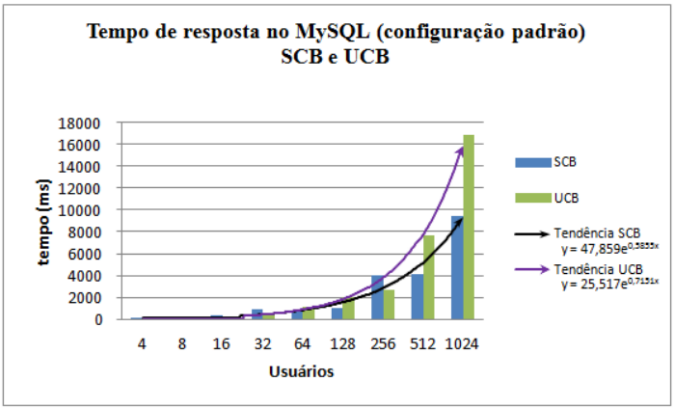


Figura 2. Tempo de resposta no MySQL com a configuração padrão utilizando *select* com carga baixa e *update* com carga baixa.

A Figura 1 e Figura 2 mostram exemplos de alguns dos testes realizados, apresentando os tempos de resposta com diferentes quantidades de usuários para MySQL e MariaDB em configuração padrão. Neste caso, o MariaDB obteve resultados mais satisfatórios, respondendo mais rapidamente as requisições que exigem pouco

esforço computacional. Enquanto que para requisições com carga maior, que exigem mais de todo o sistema, o MySQL obteve uma pequena vantagem.

Por meio dos testes realizados, observou-se que o MariaDB tem desempenho superior em termos de tempos de resposta quando exposto a situação de concorrência para simples consultas e atualizações em tabelas. Por outro lado, para uma alta carga de trabalho o MySQL se mostrou pouco superior na maioria dos testes. Em relação à utilização do recurso *GET_LOCK()*, a diferença entre os SGBDs foi mínima, embora o tempo de resposta em ambas tenha aumentado significamente. Por fim, fez se o uso do recurso *pool_of_threads* nas configurações do MariaDB, mostrando-se eficiente para quantidades iguais ou superiores a 256 usuários, melhorando o desempenho. É importante ressaltar que este recurso não foi possível ser utilizado nos testes com o MySQL, pois o mesmo é ofertado somente nas versões com pagamento de licenças comerciais.

5. Conclusão e Trabalhos Futuros

Tendo em vista os aspectos observados, conclui-se que o MariaDB pode ter um melhor desempenho em sistemas que realizem simples transações no banco de dados, mesmo que com um número elevado de usuários acessando a base de dados simultaneamente. Este desempenho pode ser aprimorado com o uso do *pool of threads* para uma quantidade maior de clientes concorrentes. Para sistemas que realizem requisições de carga de trabalho elevada, conclui-se que tanto o MariaDB e MySQL, por possuírem desempenhos bastante próximos, provavelmente terão comportamentos semelhantes.

As informações apresentadas são úteis para escolha e migração entre estes SGBDs, pois percebe-se que existem vantagens e desvantagens da utilização de cada quanto a concorrência no acesso de dados.

Como trabalhos futuros pretende-se analisar o desempenho destes SGBDs para outros tipos de transações e também verificar quais os fatores que resultam em diferenças de desempenhos. Outro trabalho pretendido é analisar a concorrência entre vários usuários realizando diversos tipos de transações ao mesmo tempo. Analisando também quais as características de SGBD influenciaram no tempo de respostas.

Também pretende-se realizar um estudo sobre outras técnicas ou recursos que contribuem na otimização dos tempo de resposta em situações de concorrência. Dessa forma, determinar quais possuem um melhor desempenho e para qual caso cada técnica será apropriada.

References

CIFERRI, Ricardo R. Um Benchmark Voltado à Análise de Desempenho de Sistemas de Informações Geográficas. 1995. 192 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Estadual de Campinas, Campinas, SP, 1995.

DAL'OSTO, Fábio. Método para Avaliação de Ambientes de Desenvolvimento de Software Combinando CMM e GQM. 2003. 110 f. Dissertação (Mestrado em Ciência da Computação) – Universidade Federal do Rio Grande do Sul, Porto Alegre, RS. 2003.

DATE, C. J. Introdução a Sistemas de Banco de Dados. 4. ed. Rio de Janeiro: Editora Elsevier, 2003.

ELMASRI, Ramez; NAVATHE, Shamakant B. Sistemas de Banco de Dados. 4. ed. São Paulo: Editora Pearson Addison Wesley, 2005.

HALILI, Emily H. Apache JMeter: A Practical Beginner's Guide to Automated Testing and Performance Measurement for your Websites. 1. ed. Birmingham: Packt publishing, 2008.

MOFFATT, Robin. Performance and OBIEE – part VI – Analysing results. Disponível em <http://www.rittmanmead.com/2013/03/performance-and-obiee-analysing-results/> Acesso em: 01 jan. 2014.

MURDOCCA, Miles J.; HEURING, Vicent P. Principles of Computer Architecture. 1. ed. Prentice Hall, 1999.

KYOUNG-DON, K.; SIN, P. H.; JISU, O., SON, S. H. A Real-Time Database Tested and Performance Evaluation. 3th IEEE International Conference. Embedded and Real-Time Computing Systems and Applications, 2007.

PERSEGUINE, Vanessa Ravazzi. Análise do controle de concorrência em diferentes tecnologias de banco de dados. 2012. Dissertação (Mestrado) - Universidade Federal de Santa Catarina, Centro Tecnológico. Programa de Pós-Graduação em Ciência da Computação. 2012.

PIRES, Carlos E. S.; NASCIMENTO, Rilson O.; SALGADO, Ana C. Comparativo de Desempenho entre Banco de dados de Código Aberto. Escola Regional de Banco de Dados. Anais da ERBD06. Porto Alegre: SBC, 2006.

RAMAKRISHNAN, Raghu; GEHRKE, Johannes. Sistemas de Gerenciamento de Banco de Dados. 3. ed. Rio de Janeiro: Editora McGraw-Hill, 2008.

SANTOS, Ismayle S.; NETO, Pedro A. S. Automação de Testes de Desempenho e Estresse com o JMeter. Livro-texto dos minicursos do II ERCEMAPI – Escola Regional de Computação, São Luiz, MA. Cap 7, p 151-176, 2008.

SCHWENKE, Axel. 5.5 Series Sysbench OLTP Results. Disponível em: <https://blog.mariadb.org/5-5-series-sysbench-oltp-results/> Acesso em: 24 jan. 2014.

SILBERSCHATZ, Abraham; KORTH, Henry F.; SUDARSHAN, S. Sistemas de Banco de Dados. 5. ed. São Paulo: Editora Elsevier, 2006.

SILVA, Ricardo C. Benchmark em Banco de Dados Multimídia: Análise de Desempenho em Recuperação de Objetos Multimídia. 2006. 64 f. Dissertação (Mestrado em Informática) – Universidade Federal do Paraná, Curitiba, PR, 2006.

SILVA, W. J.; SILVA, H. de A.; BORGES, F. F.; Desempenho do Uso de Visões nas Consultas em Diferentes Servidores de Banco de Dados Relacionais. In: EnAComp, X Encontro Anual de Computação, Universidade Federal de Goiás (Câmpus Catalão), Catalão – Goiás, Anais, 2013, 89-96.

VIEIRA, Marco.; DURÃES, João.; MADEIRA, Henrique. Especificação e Validação de Benchmarks de Confiabilidade para Sistemas Transaccionais. IEEE Latin America Transactions, Jun. 2005.

Melhoria no Processo de Software em Pequena Empresa

Alessandro Machado Dahlke, Douglas Tybusch, Gustavo Stangherlin Cantarelli

Sistemas de Informação – Centro Universitário Franciscano
97.010-032 – Santa Maria – RS – Brasil

{alessandrodahlke,douglastybusch,gus.cant}@gmail.com

Abstract. *Small businesses are most often characterized by informal processes, these often stimulated by horizontal organizational structure. In order to minimize the informality, this work aims to define an optimization proposal for the development process of a small company, thereby targeting software, the search for the same quality, as well as increased staff productivity. The proposal is based on the representation of the current tasks of the development sector and also an optimized process using BPMN notation. The optimization will be based on the ISO/IEC 12207 quality and some good practices of SCRUM methodology as needed.*

Resumo. *Pequenas empresas são, na maioria das vezes, caracterizadas pela informalidade de processos, muitas vezes estimuladas pela estrutura organizacional horizontal. A fim de minimizar tal informalidade, este trabalho tem como objetivo definir uma proposta de otimização para o processo de desenvolvimento de software de uma pequena empresa, visando assim, a busca pela qualidade do mesmo, bem como o aumento da produtividade da equipe. A proposta tem como base a representação das atuais tarefas do setor de desenvolvimento e, também, de um processo otimizado utilizando a notação BPMN. A otimização será baseada na norma de qualidade ISO/IEC 12207 e em algumas boas práticas da metodologia SCRUM conforme for necessário.*

1. Introdução

Pequenas empresas em geral, são caracterizadas por ambientes de trabalho informal e pela comunicação facilitada por estruturas organizacionais horizontais [Tolfo, Medeiros e Mombach 2013]. Tendo como foco as pequenas empresas desenvolvedoras de software, e a definição de seus processos, o presente trabalho abrange o estudo do processo atual, utilizado em uma pequena empresa, e uma proposta de otimização visando o aumento da qualidade dos produtos ofertados, bem como o aumento da produtividade da equipe.

Um software tem qualidade quando atende as necessidades do cliente e a padrões de qualidade pré-definidos [Rezende 2005]. A garantia da qualidade de software é uma atividade que é aplicada ao longo de todo o processo de engenharia de software, consiste na execução de procedimentos, técnicas e ferramentas aplicadas por profissionais para assegurar que um produto atinja padrões de qualidade pré-definidos. Estes processos devem estar bem definidos e muitas vezes não existem nas pequenas empresas. Com isso, a qualidade do processo de produção é um fator que interfere diretamente na qualidade do produto final [Koscianski 2007].

Visando a definição de processos, o objetivo deste trabalho é modelar e analisar o processo atual propondo um processo otimizado através da notação *Business Process Management Notation* (BPMN). Após a análise, a proposta será definida tendo como base a utilização das categorias e alguns sub-processos da norma ISO/IEC 12207, que define um processo de ciclo de vida de software, bem como a utilização de alguns processos da metodologia ágil SCRUM que melhor se adaptem a realidade da empresa.

2. Modelagem de processos x Pequenas empresas

A modelagem de processos é uma representação abstrata da arquitetura, projeto ou definição do processo de software. Com base nas características de pequenas empresas, tais como: comunicação simples, processos instáveis e inexperiência na área de engenharia de software, é necessário que se faça uma adaptação dos modelos existentes para que esta modelagem seja realizada [Weber, Hauck e Wangenheim 2005].

As pequenas empresas, diferentes das empresas de grande porte, na maioria das vezes não possuem recursos suficientes para destinar profissionais a desempenhar funções específicas, a qual caracteriza uma estrutura informal com pessoas que executam diferentes trabalhos conforme a necessidade. Esta estrutura permite à empresa reagir mais prontamente conforme surgem oportunidades ou problemas, mas também podem deixar confuso o processo decisório à medida que as funções se justapõem [Tolfo, Medeiros e Mombach 2013].

Slack, Chambers e Johnston (2009) verificaram que a visão por processos é válida inclusive para as empresas de pequeno porte [Tolfo, Medeiros e Mombach 2013]. Portanto, com a definição e implantação de um processo é necessário construir um modelo que o represente, o mesmo deve dar suporte ao entendimento e a visualização do processo, ajudando na avaliação, evolução e melhoria contínua do mesmo. O ideal é que cada organização não crie modelos, mas utilize com sabedoria as melhores práticas indicadas, adaptando-as conforme a realidade empresarial [Rocha, Oliveira e Bezerra 2004].

Modelar o processo dentro do ambiente em que ele será executado é bastante importante, é preciso conhecer bem a organização envolvida, conhecer as características, tipos de software que são desenvolvidos, paradigmas de desenvolvimento e cultura da empresa de forma a conhecer os processos que atendam as suas necessidades [Weber, Hauck e Wangenheim 2005].

A maioria das empresas de desenvolvimento de software nasceu pequena e desenvolveu uma cultura própria de trabalho que, em um primeiro momento, se mostrou eficaz e possibilitou o seu crescimento [Antonio, Marly e Mauro 2005]. Na maioria das vezes por estarem acostumadas com o processo de trabalho atual, que aparenta ser eficaz, acham que não há uma necessidade de investir em uma otimização devido aos seguintes fatores: o alto custo a ser investido, a falta de pessoas capacitadas e também a mudança de rotina das pessoas mais antigas na empresa, bem como a resistência as mudanças das pessoas com maior autoridade [Sommerville 2007].

3. Qualidade no processo de desenvolvimento de software

A qualidade de software compreende um processo que visa gerenciar todas as etapas e artefatos produzidos durante o ciclo de vida do software, garantindo o controle dos processos e produtos, assim como a prevenção e eliminação de defeitos. Logo é

impossível se obter um software de qualidade com processos de desenvolvimento deficientes, e com isso, é possível estabelecer duas dimensões fundamentais para a qualidade de software: a qualidade do processo e a qualidade do produto [Pressman 2007].

A qualidade de software é uma preocupação real e seus esforços têm sido realizados na busca pela qualidade dos processos envolvidos em seu desenvolvimento e manutenção. A qualidade do processo procura identificar a ausência de qualidade o quanto antes através do controle das conformidades com a especificação e correção de problemas. Para garantir a conformidade à especificação, é necessária a definição e execução das atividades durante todo o ciclo de vida do software.

Com base nisso, a ISO/IEC 12207 define uma estrutura para que uma organização defina os seus processos, abrangendo todo o ciclo de vida, desde o levantamento de requisitos, manutenção, melhora da qualidade dos processos de desenvolvimento até a retirada de uso do software [Koscianski 2007].

A estrutura dos processos da norma é classificada em três categorias:

- Fundamentais: abrange os processos: aquisição, fornecimento, desenvolvimento, operação e manutenção.
- Apoio: engloba os processos: documentação, gerência de configuração, garantia de qualidade, verificação, validação, revisão conjunta, auditoria e solução de problemas.
- Organizacionais: composta pelos processos: gerenciamento, infra-estrutura, melhoramentos e treinamento.

4. Proposta de otimização de processos

Primeiramente foi feito um comparativo entre a maneira como ocorre o atual processo de desenvolvimento de software, identificando quais os processos de metodologias existentes a empresa já estava utilizando. Em uma etapa seguinte, foram escolhidos alguns processos da metodologia SCRUM e de alguns sub-processos das categorias da norma ISO/IEC 12207 que poderiam ser utilizados.

Embora, atualmente, a empresa possua algumas boas práticas no ambiente de desenvolvimento, não há uma metodologia formalizada e padrões não são seguidos de forma adequada. Como as metodologias existentes não se encaixam por completo na realidade da empresa é necessário adaptar algumas práticas já existentes.

A empresa atualmente é composta por um diretor geral e duas equipes: uma equipe de suporte e uma equipe de desenvolvimento, ilustradas na Figura 1. A equipe de suporte é responsável pelo atendimento ao cliente, identificação de problemas, instalações do software e abertura de chamados de desenvolvimento e manutenção. Já a equipe de desenvolvimento é responsável pela implementação do software e pelos testes realizados.

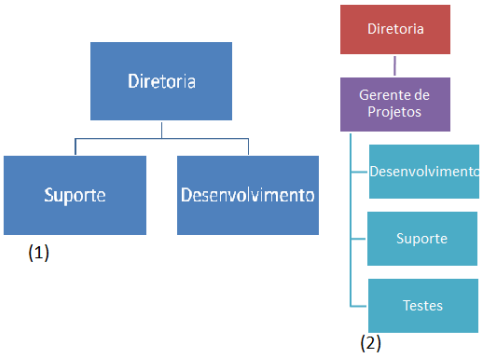


Figura 1 - Hierarquia atual (1) e proposta (2) da empresa.

As solicitações de customização do software são tratadas diretamente com o diretor da empresa, onde é marcada uma reunião para debate e análise da nova funcionalidade. Em casos de solicitação de manutenção, há um contato direto do cliente com a equipe de suporte, a qual realiza uma análise da solicitação feita e, constatando que há uma necessidade de implementação do software, é realizado o registro de um chamado para a equipe de desenvolvimento. Verificando que a solicitação trata apenas de alguma configuração ou esclarecimento de alguma dúvida, a equipe de suporte assume o papel de responsável pelo atendimento.

Atualmente, como não há um gerente de projetos para realizar o controle das atividades, a equipe de suporte acompanha as atividades em aberto através de um sistema de chamados, delegando-as conforme a identificação de quem está com o menor número de tarefas ou possui um maior conhecimento em relação ao problema.

O sistema mantém uma breve descrição do problema e a identificação de quem está atendendo-o. Após o atendimento da solicitação pela equipe de desenvolvimento, é encaminhada uma versão *beta* ao cliente onde o mesmo irá realizar testes e validar o desenvolvimento. No momento em que o cliente retorna sucesso em seus testes, a solicitação é encerrada pela equipe de suporte.

A Figura 2 ilustra o processo atual desde o contato efetuado pelo cliente até sua solicitação concluída.

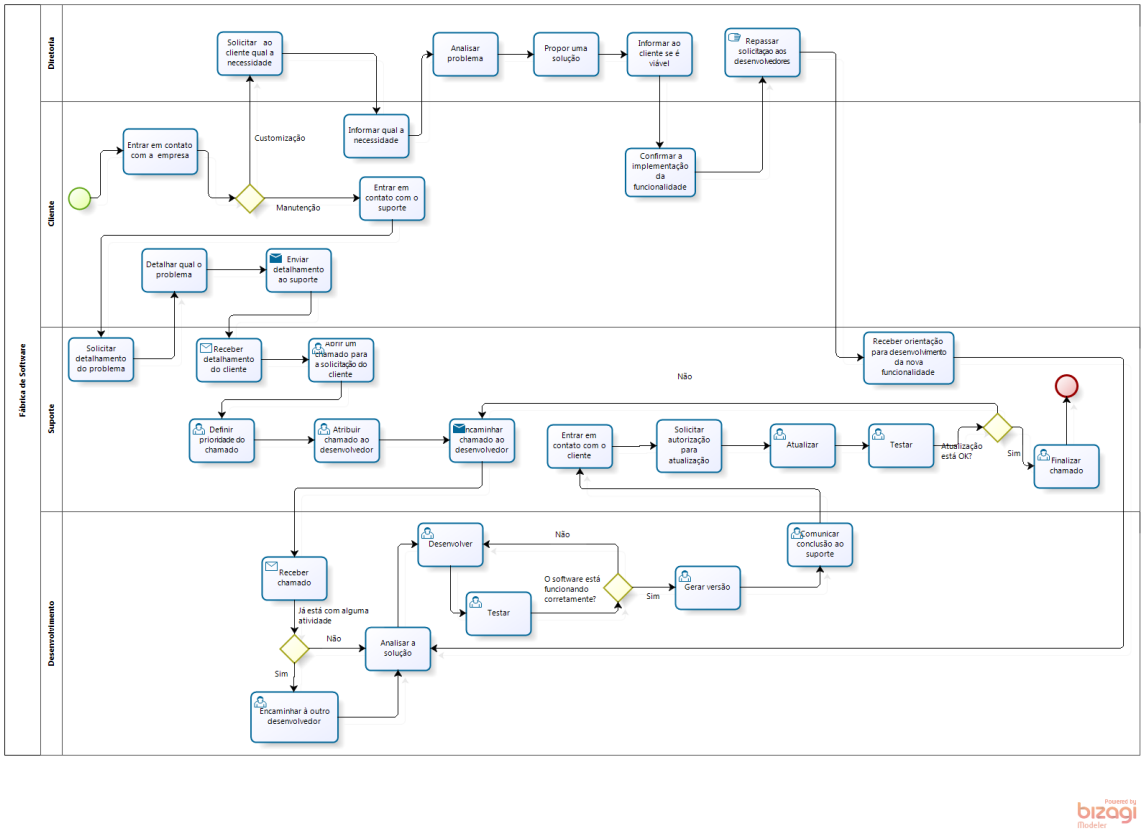


Figura 2 - Processo atual.

Como proposta de otimização do processo atual, sugere-se a utilização de ferramentas para gerenciamento de atividades, controle de versões e melhorias no fluxo do processo, definindo adequadamente os papéis desempenhados por cada uma dentro da empresa. De acordo com a norma ISO/IEC 12207 foram escolhidos alguns processos, estes mostrados na Figura 3.

Processos Fundamentais	Processos de Apoio	Processos Gerenciais
Desenvolvimento	Documentação	Treinamento
Manutenção	Gerência de Configuração	
	Garantia de Qualidade	
	Solução de Problemas	

Figura 3 - Categorias e processos utilizados conforme ISO/IEC 12207.

A categoria dos processos fundamentais contemplam os processos desde a solicitação do cliente até a entrega do software e são complementados pelos processos da categoria de apoio, principalmente com o uso de uma documentação apropriada não muito extensa. Esta, por sua vez, caracteriza-se pela elaboração do documento de requisitos, o qual irá conter a lista de requisitos funcionais e não funcionais juntamente com os diagramas de casos de uso, classes e atividades da notação UML (*Unified Modeling Language*) e, também, o histórico das alterações de cada projeto.

Conforme a categoria dos processos fundamentais sugere, no processo de desenvolvimento, os processos de análise, levantamento de requisitos, desenvolvimento, teste e implantação devem ser executados de forma correta, sem a omissão de alguma etapa. As etapas do processo de desenvolvimento devem ser executadas e revisadas sempre que houver necessidade. Para complementar este processo, serão adotados artefatos definidos na categoria dos processos de apoio. No processo de documentação, por exemplo, serão criados documentos para registro formal de todas as alterações realizadas no decorrer do projeto, tais como: documento de requisitos, para formalizar as novas funcionalidades a serem implementadas, bem como todas as modificações realizadas; documento de manutenção, para registrar todas as correções realizadas no software, incluindo correções emergenciais; e criação do documento de atualização para registrar quais funcionalidades e quais modificações está sendo disponibilizadas para cada cliente da empresa.

Além dos processos de desenvolvimento e documentação, outro processo a ser utilizado na otimização é o processo de gerência de configuração. Um profissional será responsável por disponibilizar o ambiente e manter a infraestrutura utilizada pelas equipes da empresa. As suas responsabilidades serão: manter o repositório do projeto e histórico de modificações e dar suporte à atividade de desenvolvimento dos softwares de forma que os profissionais tenham espaços de trabalho adequados para criar e testar seus desenvolvimentos.

Partindo da estrutura atual da empresa e da forma com que trabalha, o primeiro passo é reestruturar as equipes da empresa visando diminuir a carga de atividades sobre a equipe de desenvolvimento, que atualmente desenvolve e realiza os testes, a fim de atingir um aumento na qualidade dos softwares. Sugere-se a definição de uma equipe específica em testes para assumir o papel de responsável pelo planejamento e execução dos mesmos visando um melhor aproveitamento de tempo por parte das equipes e uma possível redução no número de falhas.

O processo de garantia da qualidade deverá ser executado pela equipe específica em testes de forma a minimizar as falhas existentes no software através de testes funcionais em ambiente similar ao ambiente dos clientes, com o propósito de aproximar o máximo possível à realidade do cliente.

Após a reestruturação das equipes, o foco é uma melhora na comunicação entre cliente e empresa, sugere-se que seja definido um único canal de comunicação para facilitar o esclarecimento e o entendimento da solicitação, bem como melhorar o levantamento de requisitos, mantendo um histórico com as informações trocadas em um único lugar para consulta. Cada solicitação deve ser registrada e analisada pela equipe de suporte contendo em anexo um documento de manutenção com o detalhamento do problema, facilitando o entendimento por parte da equipe de desenvolvimento e agilizando o atendimento das demais solicitações.

Propõe-se que as solicitações sejam centralizadas em uma pessoa dentro da empresa, preferencialmente que esteja mais próxima às equipes de desenvolvimento e de testes. Essa pessoa receberá o papel de gerente de projetos, que deve realizar o gerenciamento das atividades de modo que todos possam acompanhar o andamento das mesmas. Conforme a metodologia SCRUM, é proposta a utilização de reuniões diárias de no máximo 15 minutos para verificar o que já foi feito e delegar atividades em aberto aos membros da equipe. Sugere-se também a entrega frequente de funcionalidades,

tendo como padrão entregas semanais, as quais podem ser caracterizadas pela liberação de versões do software a todos os clientes.

Ao final de cada semana o gerente de projetos e o diretor da empresa devem realizar uma reunião para analisar o desempenho das equipes e também para analisar as atividades realizadas, estudar as atividades que possuem impedimentos e planejar a próxima semana. Seria interessante o gerente de projetos apresentar estes resultados para toda a equipe através da exibição de relatórios e de gráficos de desempenho, deixando todos cientes do quanto a equipe está sendo produtiva.

A Figura 4 mostra o processo otimizado e os responsáveis pela execução dos mesmos.

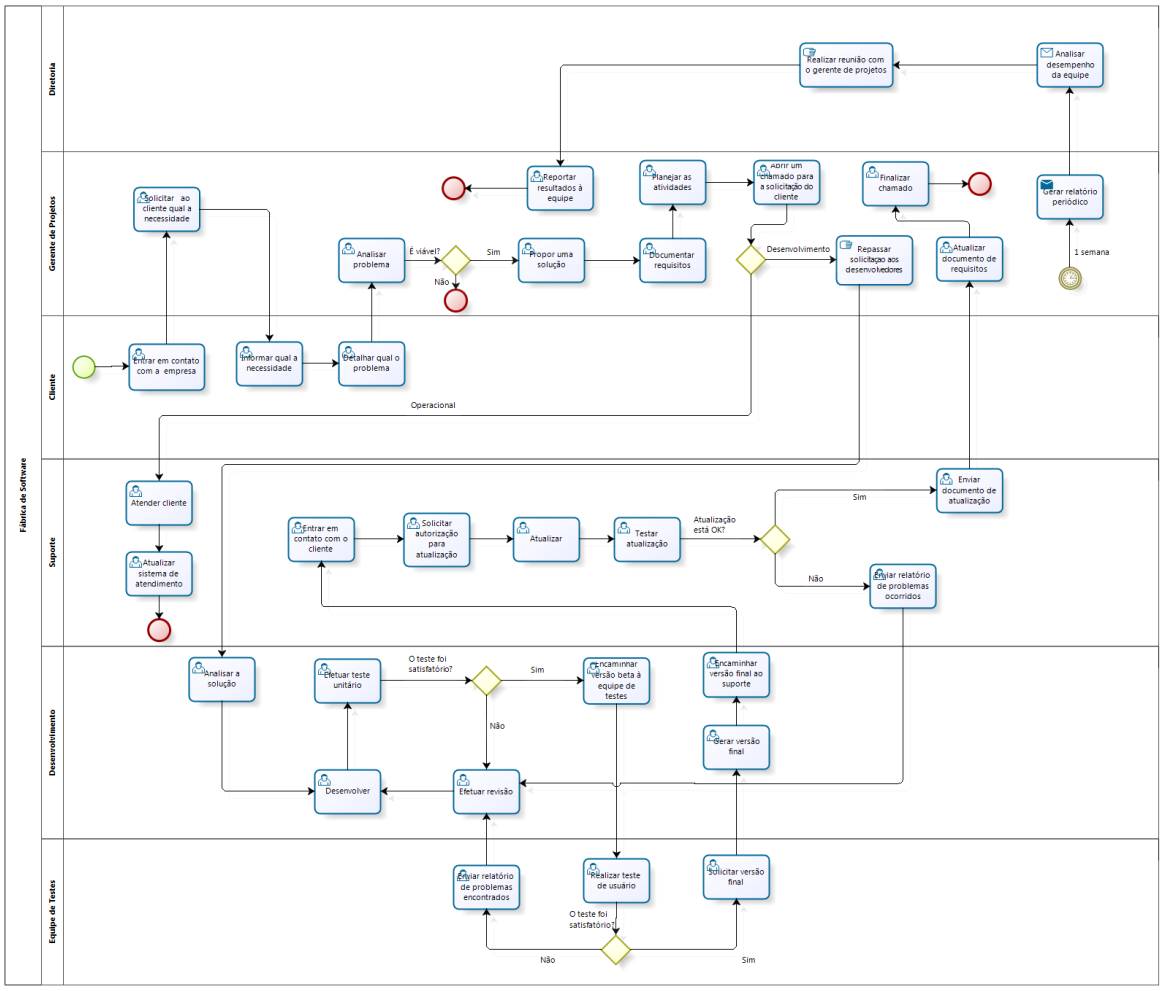


Figura 4 - Processo otimizado

5. Considerações finais

O presente trabalho teve por objetivo estudar alguns conceitos de qualidade de processos, tais como a ISO/IEC 12207 e a metodologia SCRUM e, analisar o processo atual de uma pequena empresa de desenvolvimento de software sugerindo melhorias no

mesmo. Foi analisado o processo atual desde o momento em que o cliente entra em contato com a empresa, a implementação do software até a sua implantação.

A partir da análise, verificou-se uma necessidade de melhorar o processo de software através da definição dos processos, utilização de ferramentas, bem como a definição das equipes e dos papéis desempenhados pelos envolvidos na empresa. Para definição da proposta, utilizou-se a estrutura da norma ISO/IEC 12207 e algumas boas práticas da metodologia SCRUM. Tendo em vista a cultura da empresa, pretende-se obter um aumento na produtividade da equipe, um maior controle sobre as atividades desempenhadas e a entrega de um produto final com qualidade seguindo as sugestões apresentadas neste trabalho.

6. Referências

Antonio C. T.;Marly M. C; Mauro M. S. (2005); "Integrando modelos de maturidade e qualidade – uma estratégia para a implantação de melhoria de processo de software" XXV Encontro Nac. de Eng. de Produção – Porto Alegre, RS, Brasil, 29 outubro a 01 de novembro de 2005.

Braconi, Joana; Oliveira, Saulo Barbará de. (2009) "*Business Process Modeling Notation* (BPMN)." In: Valle Rogerio; Oliveira, Saulo Barbará de. "Análise e Modelagem de Processos de Negócio: Foco na Notação BPMN." São Paulo: Atlas. p. 77-93.

Ian Sommerville: (2007) "Software Engineering", 8th edition, Pearson Education, (2007).

Koscianski A. (2007) "Qualidade de Software: aprenda as metodologias e técnicas mais modernas para o desenvolvimento de software" 2ªed. São Paulo, Nova Tec editor, (2007).

OMG. (2014) "Business Process Model & Notation (BPMN)", http://www.omg.org/bpmn/index.htm.

Rezende, D. A.; "Engenharia de Software e Sistemas de informação/Denis Alcides Rezende." 3. ed. Rio de Janeiro: Brasport, (2005).

Rocha, T. A.; Oliveira, Bezerra S.R.; Vasconcelos, Lins A. M.; (2004) "Adequação de Processos para Fábricas de Software." In: Simpósio Internacional de Melhoria de Processos de Software, 6.

Roger S. Pressman: (2007) "Software Engineering - A Practitioners approach," 7th edition, McGraw-Hill, (2007).

Slack, N.; Chamers, S.; Johnston, Robert. (2009) "Administração da Produção." 3ª ed. São Paulo: Atlas.

Tolfo C.; Medeiros; T. S; Mombach J.G; (2013) "Modelagem de Processos com BPMN em Pequenas Empresas: Um estudo de caso." XXXIII Encontro Nacional de Engenharia de Produção. Salvador, BA, Brasil, 08 a 11 de outubro de (2013).

Weber S., Hauck J.C.R., Wangenheim C.G. (2005) "Estabelecendo Processos de Software em Micro e Pequenas Empresas".

Avaliação dos recursos de acessibilidade de sistemas de compras eletrônicas. Um estudo de caso com o uso de ferramentas automáticas para avaliação de acessibilidade

Gilmara Oliveira Maquiné e Sara Valério Meirelles.

Instituto Federal de Educação, Ciência e Tecnologia do Estado do Amazonas (IFAM).

Caixa Postal 69020-120, Manaus – Amazonas - Brasil

gilmaramaquine@gmail.com, saravmeirelles@gmail.com

Abstract. *The purpose of this paper is to evaluate the degree of accessibility of two cases of Electronic Procurement Systems on the Web, focusing on initial assessment of the difficulties of use by people with visual limitations. Using automated tools for assessment and accessibility metrics proposed by the World Wide Web Consortium (W3C), which are used as guidelines for Web Content Accessibility Guidelines 1.0 (WCAG), it was possible to identify points of accessibility can be improved in these systems.*

Resumo. *A proposta deste artigo é avaliar o grau de acessibilidade de dois casos de Sistemas de Compras Eletrônicas na Web, tendo como foco inicial da avaliação as dificuldades de uso pelas pessoas com limitações visuais. Utilizando ferramentas automáticas de avaliação e as métricas de acessibilidade propostas pelo World Wide Web Consortium (W3C), que são aplicadas como Diretrizes para Acessibilidade do Conteúdo Web 1.0 (WCAG), foi possível identificar quais pontos de acessibilidade podem ser melhorados nestes sistemas.*

1. Introdução

Na atualidade com o uso popularizado da Internet, os sistemas de compras eletrônicas têm tornando a vida das pessoas mais ágil, em decorrência da facilidade na realização dos seus serviços. Estes serviços atingem um público mundialmente diversificado, e parte deste público é composto de pessoas que possuem algum tipo de deficiência.

A necessidade de desenvolver sistemas com acessibilidade para atender a todos os públicos, tem sido um desafio enfrentado pelos desenvolvedores destes sistemas em decorrência desta abordagem não ser uma maneira trivial de desenvolvimento. Ou seja, para desenvolver um sistema ou site na web que possua acessibilidade em seus conteúdos, faz-se necessária a utilização de padrões e diretrizes especificas tornando assim o desenvolvimento diferenciado do modelo tradicional.

Contudo, se os sites e sistemas em seus projetos utilizarem as métricas de acessibilidade, propostas como diretrizes pelo World Wide Web Consortium(W3C), as pessoas que possuem algum tipo de necessidade especial poderão fazer uso dos sites e sistemas em sua completude.

Para tanto, propõe-se uma avaliação de dois Sistemas de Compras Eletrônicas utilizando duas ferramentas de avaliação automáticas, que são baseadas nas métricas de

acessibilidade propostas pelo W3C, de modo que possam ser identificadas melhorias para tornar os sistemas mais acessíveis à todos os públicos.

2. Acessibilidade Web

O tema acessibilidade vem sendo destacado nas diversas áreas de conhecimento, em decorrência da necessidade da utilização de serviços disponíveis na Web. GAMELEIRA, na Cartilha para Acessibilidade, conceitua acessibilidade da seguinte maneira: “Acessibilidade é a tradução operacional do direito básico de ir e vir, de forma independente, em todos os ambientes sejam físicos ou virtuais.”.

Para NICHOLL, acessibilidade é a possibilidade de qualquer pessoa independentemente de suas capacidades físico – motoras e perceptivas, culturais e sociais, usufruir os benefícios de uma vida em sociedade, ou seja, é a possibilidade de participar de todas as atividades, até as que incluem o uso de produtos serviços e informações com o mínimo de restrições possível.

Na Acessibilidade na Web é um conceito que vem evidenciando a necessidade de aplicar padrões ao desenvolvimento de sistemas web, de modo que possibilite a utilização dos sistemas por todos os tipos de públicos.

Segundo FREITAS, acessibilidade na Web corresponde a possibilitar que qualquer usuário, utilizando qualquer agente (software ou hardware que recupera e serializa conteúdo Web) possa entender e interagir com o conteúdo do sítio.

CONFORTO destaca a relevância da discussão da Acessibilidade à Web da seguinte maneira:

> Discutir a temática da Acessibilidade Web é impulsionar a concretização dos quatro movimentos necessários para a construção de uma sociedade potencialmente inclusiva e democrática:
>
> Qualidade de Vida – democratizar os acessos às condições de preservação e de desenvolvimento do homem e do meio ambiente;
>
> Autonomia – capacitar sujeitos a supriram suas necessidades vitais, culturais e sociais;
>
> Desenvolvimento Humano – possibilitar o desenvolvimento de capacidades intelectuais e biológicas;
>
> Equidade – garantir a igualdade de direitos e oportunidades, respeitando as especificidades da diversidade humana.

As Diretrizes para Acessibilidade do Conteúdo Web 1.0 *(Web Content Accessibility Guidelines 1.0 - WCAC),* foram criadas pelo W3C com o objetivo de explicar como criar conteúdos na Web acessível à pessoas com deficiência. As orientações contidas na WCAG são destinadas a todos os desenvolvedores de conteúdo web. No total são 14 diretrizes que foram criadas pela W3C, e encontram-se listadas na Tabela 1.

Tabela 1 - Diretrizes para Acessibilidade do Conteúdo Web 1.0 (*WCGA 1.0*) Fonte:W3C

Diretiva	**Título**	**Objetivo**
1	Fornecer alternativas ao conteúdo sonoro e visual.	Proporcionar conteúdo que, ao ser apresentado ao usuário, transmite essencialmente a mesma função e finalidade que o conteúdo sonoro ou visual.
2	Não recorrer apenas a cor.	Garantir que o texto e os gráficos são compreensíveis quando vistos sem cores.
3	Utilizar corretamente marcações e folhas de estilo.	Marcar os documentos com os elementos estruturais adequados. Controlar a apresentação com folhas de estilo e não com elementos de apresentação e atributos.
4	Indicar claramente qual o idioma utilizado	Use marcação que facilitem a pronúncia e a interpretação de texto abreviado ou estrangeira.
5	Criar tabelas passíveis de transformação harmoniosa.	Assegurar que as tabelas possuam a marcação necessária para serem transformadas por navegadores acessíveis e outros agentes de utilizador.
6	Assegurar que as páginas dotadas de novas tecnologias sejam transformadas harmoniosamente.	Assegurar que as páginas são acessíveis mesmo quando as novas tecnologias não são suportados ou estão desligados.
7	Assegurar o controle do usuário sobre as alterações temporais do conteúdo	Certifique-se que em movimento, deslocamento, piscando, ou actualização automática de objectos ou páginas podem ser interrompidas ou paradas.
8	Assegurar a acessibilidade direta de interfaces do usuário integradas	Certifique-se que a interface do usuário segue os princípios de desenho acessível: acesso independente do dispositivo de funcionalidade, operacionalidade pelo teclado, auto-expressar, etc
9	Projetar páginas considerando a independência de dispositivos	Use os recursos que permitam a ativação de elementos de página através de uma variedade de dispositivos de entrada.
10	Utilizar soluções de transição	Utilizar soluções de acessibilidade transitórias, de modo que as tecnologias de apoio e os navegadores mais antigos funcionem corretamente.
11	Utilizar tecnologias e recomendações do W3C	Use tecnologias W3C (de acordo com as especificações) e seguir as diretrizes de acessibilidade. Sempre que não seja possível a utilização de uma tecnologia W3C, ou fazendo resultados assim no material que não se transforma graciosamente, fornecer uma versão alternativa do conteúdo que está acessível.
12	Fornecer informações de	Fornecer informações de contexto e orientação para ajudar os

	contexto e orientações	usuários a compreenderem páginas ou elementos complexos.
13	Fornecer mecanismos de navegação claros	Fornecer mecanismos de navegação claros e coerentes - informações de orientação, barras de navegação, um mapa do site, etc - para aumentar a probabilidade de que uma pessoa vai encontrar o que eles estão procurando em um site.
14	Assegurar a clareza e simplicidade dos documentos	Assegurar que os documentos são claros e simples para que eles possam ser mais facilmente compreendido.

Cada uma destas diretrizes possui pontos de checagem, ou checkpoints, que explicam como cada diretriz é aplicada em cenários típicos de desenvolvimento.

3. Estudo de caso

Os sistemas de compras eletrônicas vêm se tornando um serviço bastante utilizado por pessoas de todo o mundo, e por este motivo devem estar acessíveis e prontos para anteder tanto ao público geral quanto aos que possuem alguma dificuldade ou acessam de maneira diferente.

O comércio eletrônico (*e-commerce*) é definido por LUCIANO e FREITAS, como o compartilhamento de informações de negócio, manutenção de relações de negócios e condução de transações por meio de redes de telecomunicação.

Foram escolhidos dois sistemas de compras eletrônicas para serem avaliados. Por questões acadêmicas serão ocultos os nomes dos sistemas, sendo disponibilizadas apenas imagens da homepage dos sistemas.

Figura 1 – Homepage do Sistema A

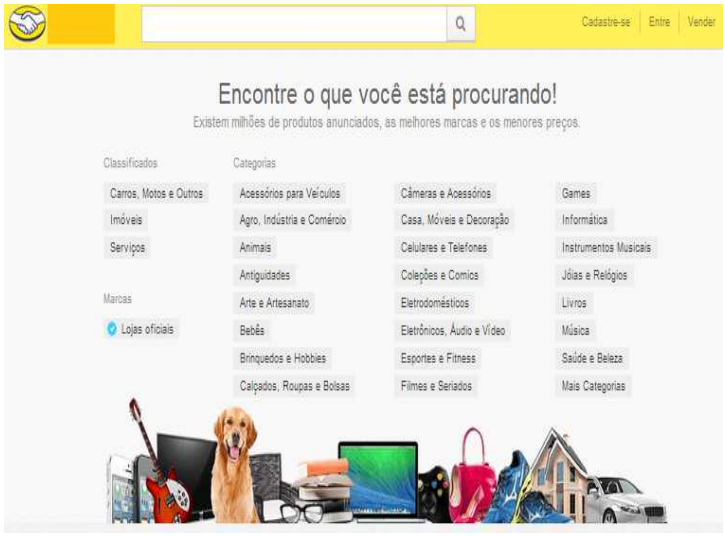


Figura 2 – Homepage do Sistema B

3.1 As ferramentas utilizadas na avaliação

Para avaliar os Sistemas de Compras Eletrônicas elencamos as ferramentas descritas na Tabela 1, seguindo os critérios de eleição também descritos na Tabela 2.

Tabela 2 - Critérios de eleição das ferramentas de avaliação

Critério	Peso	Hera 2.1 Beta	Total Hera	Cyntia Says	Total Cyntia	Examinator	Total Examinator
Facilidade de uso	3	3	9	1	3	3	9
Linguagem em Português	2	2	4	0	0	2	4
Quantidade de Referências	2	1	2	1	2	2	1
Soma Geral	-	-	**15**	-	5	-	**15**

Após analise seguindo os critérios elencados na Tabela 2, foram eleitas as ferramentas Hera 2.1 Beta e Examinator.

3.1.1 Hera 2.1 Beta

O HERA é uma ferramenta semi-automática que rever a acessibilidade das páginas Web de acordo com as recomendações das *Diretrizes de Acessibilidade para o Conteúdo Web 1.0* (WCAG 1.0). Ela facilita o trabalho do revisor, indicando os pontos que com segurança estão em falta, os que com segurança se encontram bem, os que não se aplicam na página e os pontos que obrigatoriamente devem ser revistos por mãos humanas para comprovar realmente que a página é acessível.

O HERA foi desenhado e desenvolvido no início do ano de 2003 por Carlos Benavídez, com a colaboração de Emmanuelle Guiteérrez y Restrepo e Charles McCathieNevile especialmente para a **Fundación Sidar**. O nome Hera que além de ser de uma personagem feminina mitológica corresponde o acrônimo de Folhas de Estilo para a Revisão da Acessibilidade.

A ferramenta permite localizar erros através de um visão da página, apresentando os elementos que necessitam de revisão utilizando para melhor destaque os ícones, moldura de cor, e muitas vezes os códigos dos elementos.

3.1.2 Examinator

A segunda ferramenta destacada nos critérios de eleição descritos no item 3.2 foi o Examinator. Esta ferramenta é definida como um validador automático do grau de satisfação, por uma dada página na Internet, das Diretrizes de Acessibilidade para o Conteúdo da Web (*WCAG 1.0*). Segundo informações escritas na nota técnica sobre o validador Examinator, ele é uma das 130 validadores automáticos existentes.

Este validador foi criado pela *UMIC* – Agência para a Sociedade do Conhecimento, IP com o objetivo de ultrapassar as limitações apresentadas por outros validadores e tem como vantagem a possibilidade de verificar todas as páginas do sítio.

Se comparado com outros validadores o Examinator destaca-se pelo fato de não ser necessária a instalação do software para ter acesso ao serviço. Desta maneira, para fazer a validação de qualquer site ou página da Internet faz-se necessário apenas acessar o link onde encontra-se hospedado o serviço.

O Examinator possui pelo menos 44 pontos de verificação das *WCAG 1.0*, com pelo menos um teste para cada um destes. Na Figura 5 podem ser conferidos a quantidade de testes disponíveis por ponto de verificação agrupados por nível de prioridade.

3.2 Resultados analisados

O Sistema A foi analisado com as duas ferramentas: HERA e *Examinator*, apresentando os seguintes resultados:

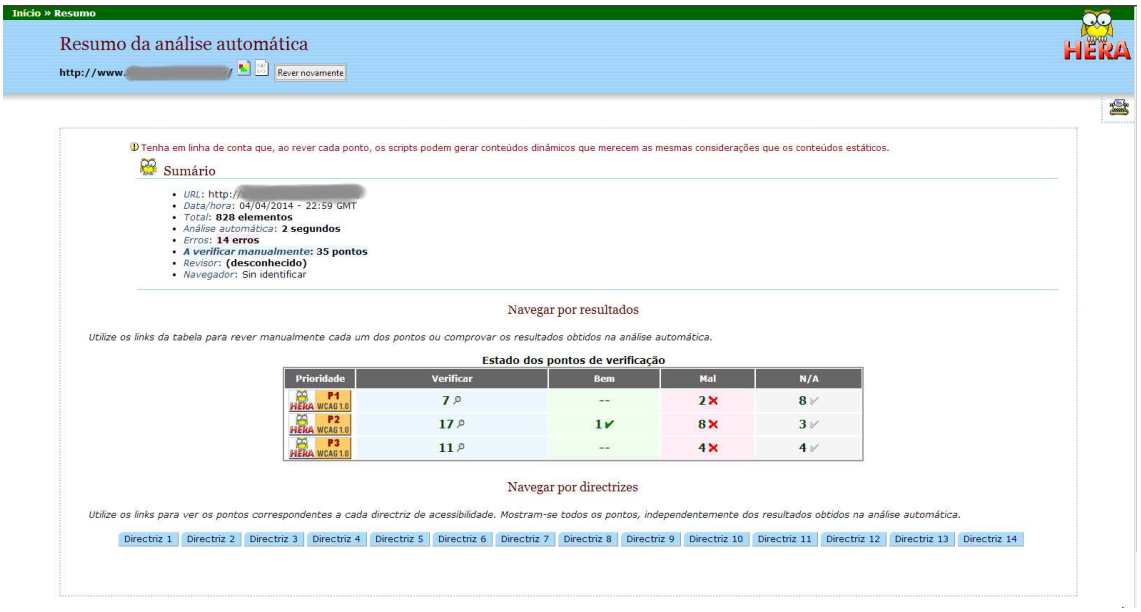


Figura 4 – Resultados de Testes com o Sistema A utilizando a ferramenta Hera.

Figura 5 – Resultados de Testes com o Sistema A utilizando a ferramenta Examinator.

O Sistema B também foi analisado com as ferramentas Hera e *Examinator* apresentando os seguintes resultados:

Figura 6 – Resultados de Testes realizados com o Sistema B utilizando a ferramenta Hera.

Figura 7 – Resultados de Testes realizados com o Sistema B utilizando a ferramenta Examinator.

4. Considerações finais

Com as informações expostas neste artigo torna-se ainda mais evidente que a acessibilidade na Web é um tema que deve ser levado em consideração desde projeto de

desenvolvimento dos sites ou sistemas que disponibilizarão serviços ao público em geral. Existe a vantagem das *Diretrizes de Acessibilidade para o Conteúdo Web* (*WCGA 1.0)*, que possibilitam realizar o desenvolvimento de sistemas web tornando o conteúdo completamente acessível para todos os públicos.

A avaliação proposta tem finalidade acadêmica e poderá ser aprofundada em trabalhos futuros com o detalhamento dos resultados para apontar onde os sites podem ser melhorados.

Os detalhes das avaliações podem ser conferidos no site: https://sites.google.com/site/avaliacaodeacessibilidadenaweb/

5. Referências

CONFORTO, D. E. A. (2010) “Tecnologias digitais acessíveis”. Porto Alegre: JSM Comunicação LTDA.

FREIRE, A. P. (2008) “Acessibilidade no desenvolvimento de sistemas web: um estudo sobre o cenário brasileiro”. [S.l.]: [s.n.]. Dissertação de Mestrado em Ciências - Ciência da Computação e Matemática Computacional.

LUCIANO, Edimara Mezzomo. e FREITAS, Henrique Mello Rodrigues. “Comércio Eletrônico de Produtos Virtuais: definição de um Modelo de Negócios para a comercialização de software”.

QUEIROZ, M. A. (2008) “Métodos e Validadores de Acessibilidade Web”. http://acessibilidadelegal.com/13-validacao.php, Junho.

“Web Content Accessibility Guidelines 1.0”. http://www.w3.org/TR/WAI-WEBCONTENT/, Junho.

GAMELEIRA, F.A.B. “Cartilha de Acessibilidade”. http://www.lupadigital.info/, Junho.

EXAMINATOR. “Nota técnica sobre o validador eXaminator (versão WCAG 1.0)”. http://www.acessibilidade.gov.pt/webax/nota_tecnica.html, junho.

HERA 2.1 BETA. “Revendo a Acessibilidade com Estilo”. http://www.sidar.org/hera/ , junho.

Árvores de Decisão Aplicadas na Previsão de Desempenho de Alunos: Estado da Arte

Marcos Alves Vieira[1], **Ernesto Fonseca Veiga**[1]

[1]Instituto de Informática – Universidade Federal de Goiás (UFG)
Goiânia – GO – Brasil

{marcosalves, ernestofonseca}@inf.ufg.br

Abstract. *Educational Data Mining (EDM) is a field that uses machine learning, data mining, and statistics to process educational data, aiming to reveal useful information for analysis and decision making. This work aims to present a survey of the state of the art through the analysis of relevant and recent articles in the field, focusing on methods that use decision trees for predicting students performance. Additionally, it presents the concepts and basic foundations of EDM.*

Resumo. *A mineração de dados em ambientes educacionais (EDM) é uma área que faz uso de aprendizagem de máquina, mineração de dados e estatística para processar dados educacionais, com objetivo de revelar informações úteis para análise e tomada de decisão. Este trabalho tem a finalidade de apresentar um levantamento do estado da arte, por meio da análise de artigos relevantes e recentes na área, com foco nos que utilizam métodos de árvores de decisão para a previsão de desempenho de alunos. Além disso, também são apresentados os conceitos e fundamentos de EDM.*

1. Introdução

A mineração de dados em ambientes educacionais (*Educational Data Mining* - EDM), é uma área emergente que se vale de técnicas de mineração de dados para revelar informações importantes, presentes em bases de dados de instituições de ensino que, de outra forma, seriam dificilmente perceptíveis. Esta área ganha ainda mais importância nos dias de hoje com a crescente utilização de plataformas de aprendizagem informatizadas, os chamados sistemas de gestão de aprendizagem (*Learning Management Systems* - LMS), inclusive em se tratando de grandes volumes de dados.

A utilização da EDM permite descobrir novos conhecimentos com base em dados gerados pelos alunos, de modo a ajudar a validar e avaliar os sistemas de ensino, visando melhorar alguns aspectos da qualidade da educação, além de estabelecer bases para um processo de aprendizagem mais eficaz. Há ainda outros exemplos de utilização de EDM, tais como, guiar o aprendizado dos estudantes e obter métodos que possam avaliar uma nova metodologia de ensino [Romero and Ventura 2010].

O restante desse trabalho está estruturado como se segue. A Seção 2 aborda os fundamentos básicos de EDM. Na Seção 3, é realizado um levantamento do estado da arte em EDM, analisando trabalhos recentes de maior relevância nesta área. E, finalmente, na Seção 4, são apresentadas as conclusões deste trabalho.

2. Mineração de Dados Educacionais

A quantidade de dados armazenados em bancos de dados educacionais aumenta rapidamente. À medida em que isto acontece, diferentes técnicas de mineração de dados têm sido desenvolvidas e utilizadas com a finalidade de obter informações a partir de tais dados e também para encontrar relações ocultas entre as variáveis utilizadas [Han et al. 2006].

A mineração de dados (*data mining*) é uma ampla área que integra diferentes técnicas, incluindo aprendizado de máquina, estatística, reconhecimento de padrões, inteligência artificial e sistemas de bancos de dados para a análise de grandes volumes de informações [Wu et al. 2008]. Assim sendo, mineração de dados é um processo para extração de padrões previamente desconhecidos, válidos e potencialmente úteis que estão escondidos em grandes conjuntos de dados, conduzindo a um nível incremental de informação e conhecimento [Connolly and Begg 2005, Aviad and Roy 2011].

Por sua vez, a mineração de dados em ambientes educacionais (EDM) é um campo que explora estatística, aprendizado de máquina e algoritmos de mineração de dados aplicados a diferentes tipos de dados de ensino. Seu principal objetivo é analisar estes dados a fim de resolver questões de investigação educacional. A EDM está preocupada com o desenvolvimento de métodos para explorar os dados presentes em ambientes de ensino e, através destes métodos, compreender melhor os estudantes e as condições em que eles aprendem [Baker et al. 2010].

A grande quantidade de dados gerada pode ser atribuída ao crescimento de dois tipos de ferramentas voltadas para o ensino. Por um lado, o aumento tanto de *softwares* educacionais, como dos bancos de dados de informação sobre estudantes, criaram grandes repositórios de dados que refletem em como os alunos aprendem [Koedinger et al. 2008]. Por outro lado, o uso da Internet na educação criou um novo contexto de educação, conhecido como *e-learning*, em que grandes quantidades de informação sobre a interação ensino-aprendizagem são constantemente geradas e disponibilizadas de forma ubíqua [Castro et al. 2007].

O processo de EDM converte os dados brutos, provenientes dos sistemas de ensino, em informações úteis que podem ter um grande impacto na pesquisa e na prática educativa. Este processo não difere muito de outras áreas de aplicação de técnicas de mineração, como negócios, genética, medicina, entre outras, porque segue os mesmos passos do processo geral de mineração de dados [Romero et al. 2004]: pré-processamento, mineração de dados, e pós-processamento.

Do ponto de vista prático, a EDM permite, por exemplo, a descoberta de novos conhecimentos com base em dados de uso dos alunos, a fim de ajudar a validar e avaliar os sistemas educacionais e melhorar potencialmente alguns aspectos da qualidade da educação, além de estabelecer as bases para um processo de aprendizagem mais eficaz [Romero et al. 2004]. Algumas ideias semelhantes já foram utilizadas com sucesso em sistemas *e-commerce*, a primeira e mais popular aplicação de mineração de dados [Raghavan 2005], para determinar os interesses dos clientes e aumentar as vendas.

A EDM envolve diferentes grupos de usuários ou participantes. Diferentes grupos, olham para informações educacionais a partir de ângulos diferentes, de acordo com sua responsabilidade, visão e objetivos para a utilização da mineração dos dados educacionais [Hanna 2004]. Por exemplo, o conhecimento descoberto por algoritmos de EDM

pode ser usado não só para ajudar os professores na gestão de suas aulas, a compreender os processos de aprendizagem de seus alunos e refletir sobre os seus próprios métodos de ensino, mas também para apoiar as reflexões sobre a situação dos alunos e dar um *feedback* aos mesmos [Merceron and Yacef 2005].

Atualmente, há uma grande variedade de sistemas e ambientes de ensino, tais como: a sala de aula tradicional, *e-learning*, LMS, Sistemas Hipermídia Adaptativos, testes/questionários, textos/conteúdos. Outras ferramentas também têm se destacado e contribuído nesse contexto, como por exemplo: objetos de aprendizagem, repositórios, mapas conceituais, redes sociais, fóruns, ambientes de jogos educativos, ambientes virtuais, ambientes de computação ubíqua, etc. Cada um desses sistemas e ambientes educacionais mencionados anteriormente fornece diferentes tipo de dados, permitindo assim, que diversos problemas e tarefas sejam resolvidos através da utilização de técnicas de mineração [Romero and Ventura 2010].

3. Aplicação de Árvores de Decisão em Mineração de Dados Educacionais

Romero e Ventura [Romero and Ventura 2010] mostraram que árvores de decisão é uma abordagem útil em três áreas de EDM. São elas:

1. **Fornecimento de *feedback* para apoiar professores**: tem como objetivo fornecer *feedback* para apoiar os coordenadores de curso, professores e administradores na tomada de decisão, para, por exemplo, melhorar a aprendizagem dos alunos e organizar recursos institucionais de forma mais eficiente. Além disso, permite que os coordenadores tomem medidas proativas apropriadas, apoiadas em dados, e/ou medidas corretivas. Muitos estudos se aplicam a comparar vários modelos de mineração de dados que fornecem *feedback* que podem ser úteis nos objetivos citados, como: regras de associação, *clustering*, classificação e análise de padrões sequenciais.
2. **Recomendações para estudantes**: o objetivo das aplicações nesta área é ser capaz de fazer recomendações diretamente aos alunos no que diz respeito às suas atividades pessoais, *links* para visitas, a próxima tarefa ou exercício a ser feito, entre outras. Outros objetivos são a capacidade de se adaptar a conteúdos de aprendizagem, interfaces e ao progresso de cada aluno em particular. Diversas técnicas de mineração de dados têm sido utilizados para esta tarefa, onde as mais comuns são: de mineração de regras de associação, *clustering*, e mineração de padrões sequenciais. Estas técnicas visam identificar as relações entre as ocorrências de eventos para descobrir se existe alguma ordem específica.
3. **Previsão de desempenho de alunos**: tem como objetivo estimar o valor desconhecido de uma variável que descreve o estudante, de forma particular, voltando-se para o seu desempenho. A previsão de desempenho de alunos é uma das mais antigas e mais populares aplicações de EDM. Diferentes tipos de modelos de redes neurais têm sido utilizados para previsão de notas de estudantes; previsão do número de erros que um aluno cometerá em uma avaliação; e previsão do resultado final de alunos em disciplinas (aprovado ou reprovado).

A seguir é apresentado um levantamento de trabalhos recentes e relevantes em mineração de dados educacionais com emprego de árvores de decisão para previsão de desempenho de alunos.

Tabela 1. Descrição dos atributos e seus possíveis valores

Attribute	Description	Possible Values
Branch	Student's branch in B.Tech course.	CS, IT, EC, EN
10^{th}%	Percentage of marks obtained in 10^{th} class examination.	First $>$ 60%, Second $>$ 45 & $<$ 60%, Third $>$ 35 & $<$ 45%
12^{th}%	Percentage of marks obtained in 12^{th} class examination.	First $>$ 60%, Second $>$ 50 & $<$ 60%
B. Tech	Percentage of marks obtained in B.Tech course.	First $>$ 60%, Second $>$ 50 & $<$ 60%
Final Grade	Final Grade obtained after analysis the passing percentage of 10^{th}, 12^{th}, B.Tech.	Excellent, Good, Average

3.1. Previsão de Desempenho de Alunos

O monitoramento de desempenho envolve avaliações que possuem um papel vital no fornecimento de informações, que é voltado para ajudar os alunos, professores e administradores na tomada de decisões [Pellegrino et al. 2001]. Os fatores de mudança na educação contemporânea têm levado à busca de eficácia e eficiência no monitoramento do desempenho dos alunos em instituições de ensino, que agora está se afastando das técnicas tradicionais de medição e avaliação, para o uso da mineração de dados, que emprega várias técnicas e métodos de investigação para isolar informações importantes que estão implícitas ou ocultas.

O desempenho dos alunos em cursos universitários é de grande importância para o ensino superior, onde vários fatores podem influenciar nos resultados obtidos [Osmanbegović and Suljić 2012]. Em uma universidade, o desempenho geral de um aluno é determinado pelos resultados de avaliações realizadas. Tais avaliações podem ser de diferentes tipos, como desempenho em sala de aula, provas, questionários, trabalhos de laboratório, monitorias, envolvimento em atividades adicionais, currículo, etc.

No objetivo de avaliar e prever o desempenho de estudantes, processos de mineração de dados, especialmente a classificação, vêm sendo aplicados para ajudar no reforço da qualidade do sistema de ensino superior por meio da avaliação dos estudantes. Singh e Kumar [Singh and Kumar 2012] utilizam o método de classificação por árvores de decisão para gerar regras a partir de uma seleção de atributos que podem influenciar o desempenho dos alunos do curso de Bacharelado em Tecnologia do *Bansal Institute of Engineering & Technology*. Estas regras são posteriormente estudadas e avaliadas, e através de um sistema que as utiliza, é possível prever a proporção de estudantes que falham ou são aprovados no exame final.

O primeiro passo para a classificação, neste caso, foi determinar os grupos em que os alunos poderiam se encaixar. Para isso, foram determinados três grupos de acordo com o desempenho: *médio*, com desempenho entre 35 (inclusive) e 45%; *bom*, com desempenho entre 45 (inclusive) e 60%; e *excelente*, com desempenho maior ou igual a 60%. O segundo passo foi determinar os atributos mais significativos, que estão listados na Tabela 1, onde também é feita uma descrição dos mesmos e citados quais valores estes atributos podem assumir. A partir da base de dados dos alunos é montada a árvore de decisão, em que cada nó ramo vai representar uma escolha entre as possíveis alternativas, e cada folha representará uma decisão.

A partir desta árvore de decisão, as regras foram extraídas e aplicadas aos dados de um conjunto de 40 alunos, de quatro "ramos" distintos do curso de Bacharelado em Tecnologia, que compõem a Tabela 2. A Tabela 3 apresenta o resultado da classificação.

Tabela 2. Siglas e nomes das áreas do curso de Bacharelado em Tecnologia

Branch Code	Name of Branches
CS	Computer Science & Engineering
IT	Information Technology
EC	Electronics & Communication Engineering
EN	Electrical & Electronics Engineering

Tabela 3. Classificação final dos estudantes

Branch	No. Students	No. Students Excellent	No. Students Good	No. Students Average
CS	10	6	3	1
IT	10	5	3	2
EC	10	4	4	2
EN	10	5	3	2
Total	40	20	13	7

Objetivando melhorar o desempenho geral de uma instituição, as performances individuais devem ser examinadas. Por isso, é útil para as instituições de ensino analisar o desempenho de seus alunos para identificar as áreas de fraqueza, orientando seus alunos para um futuro melhor. Khatwani e Arya [Khatwani and Arya 2013] propõem um algoritmo baseado em árvores de decisão e algoritmos genéticos, para prever o desempenho de um aluno. O algoritmo ID3 é utilizado para criar várias árvores de decisão, cada uma prevendo o desempenho de um aluno com base em um conjunto de características diferentes. Uma vez que cada árvore de decisão fornece uma visão para o desempenho provável de cada estudante e diferentes árvores dão resultados diferentes, não é possível prever o desempenho, mas é possível identificar as áreas ou características que são responsáveis pelo resultado previsto. Para maior precisão, é incorporado um algoritmo genético que realiza iterações evolutivas para refinar os resultados.

Em [Romero et al. 2013], os autores propõem uma ferramenta de mineração de dados educacionais integrada no ambiente educacional, de maneira que todos os processos de mineração de dados possam ser realizados pelo próprio instrutor, em uma aplicação única, possibilitando que os resultados obtidos sejam aplicados diretamente no ambiente educacional. Para coletar os dados dos usuários, foi desenvolvido um *Moodle* específico, que foi aplicado em diversas universidades, sendo a primeira a Universidade de Córdoba. Os dados coletados são tratados com técnicas de árvores de decisão, redes neurais e lógica *fuzzy*, a fim de prever as notas que os estudantes universitários irão obter no exame final de seu curso.

Um problema recorrente é o da grande quantidade de dados acumulados nas bases do sistema de administração educacional, que são carentes de ferramentas de análise inteligentes. Long e Wu [Long and Wu 2012] utilizaram o algoritmo de árvore de decisão C4.5 [Quinlan 1993] para construir um modelo de análise de desempenho dos alunos que tem como objetivo fornecer as informações necessárias para a melhora da qualidade do ensino e até mesmo para tomada de decisões em futuras reformas no sistema atual.

Na Figura 1, é mostrado o fluxograma do procedimento adotado em [Long and Wu 2012] para o desenvolvimento do modelo de classificação. A primeira

Tabela 4. Atributos e valores que podem ser assumidos

Attribute name	Code	Attribute Value
Total result	A	c1, c2, c3 (excellent, good, poor)
Experimental result	B	c1, c2, c3 (excellent, good, poor)
Paper diff	C	d1, d2, d3 (high, medium, low)
Teacher eval	D	c1, c2, c3 (excellent, good, poor)
Interest	E	y, n (interested, uninterested)

parte realizada no processo de mineração foi a integração dos dados de várias bases, em uma base unificada. Como alguns campos de dados podem ser considerados como atributos de baixa relevância ou redundantes para a classificação do desempenho, o segundo passo é um processo de redução e generalização de atributos, que reflete diretamente na redução da sobrecarga computacional desnecessária.

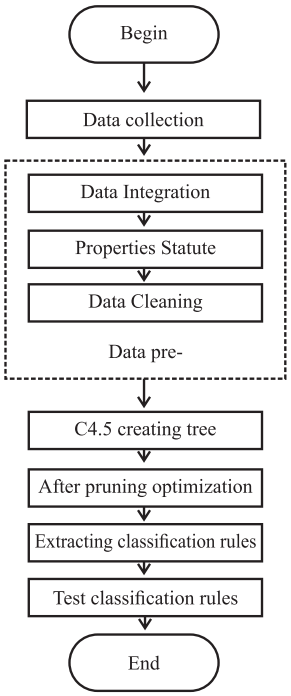


Figura 1. Fluxograma da abordagem proposta em [Long and Wu 2012].

Após a unificação da base de dados e da remoção dos dados que não possuem importância para a classificação, Long e Wu [Long and Wu 2012] aplicam um processo de limpeza dos dados (*data cleaning*) para organizar os dados que foram generalizados. A Tabela 4 mostra os atributos significativos para o problema e os possíveis valores que estes podem assumir.

Na construção da árvore de decisão, os autores empregam o algoritmo C4.5 [Quinlan 1993]. Para a escolha do melhor atributo, são realizados diversos cálculos como a entropia e o ganho de informação. O atributo com maior ganho de informação, por exemplo *B* (Figura 2) se torna o nó raiz da análise de desempenho dos alunos. De acordo com os possíveis valores de *B*, são criados diferentes ramos. O processo é repetido com os demais atributos até gerar a árvore de decisão que é mostrada na Figura 2 (esquerda).

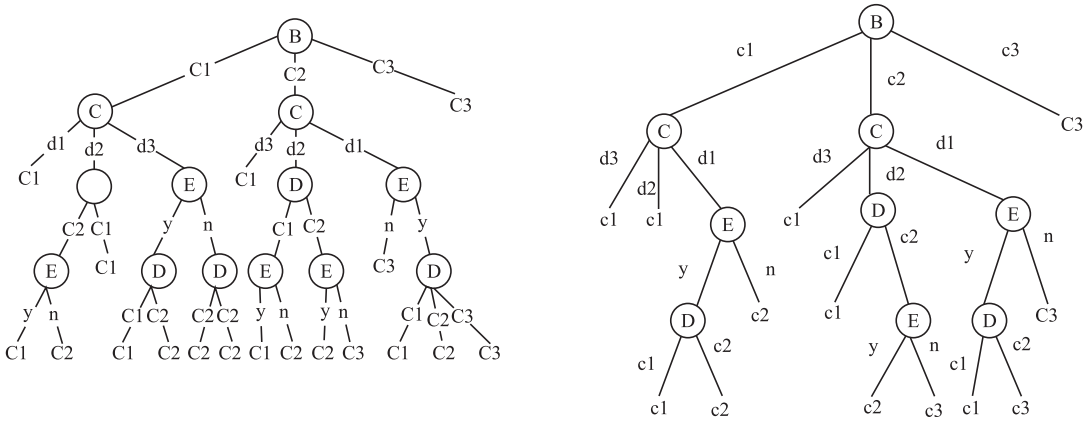


Figura 2. Árvore de decisão gerada a partir dos atributos, antes e depois do processo de poda.

No processo de criação de uma árvore de decisão, por causa dos possíveis ruídos presentes em algumas das amostras, bem como os possíveis valores incorretos no conjunto de treinamento, podem ocorrer anomalias em alguma ramificação da árvore. A fim de melhorar os resultados da árvore gerada e também reduzir sua complexidade, é utilizado um processo de poda. A Figura 2 (direita) resulta da poda da árvore da Figura 2 (esquerda). Após o processo de poda, as regras de classificação são extraídas e utilizadas para classificar as amostras da base de dados.

4. Conclusões

As metodologias de aprendizagem estão ultrapassando os limites dos ambientes das salas de aula tradicionais, contribuindo desta forma para o surgimento de ambientes dinâmicos e informatizados. Uma das consequências oriundas desta migração é o crescimento contínuo dos volumes dos dados armazenados em bases educacionais. Isto, aliado à atual mudança nas metodologias de ensino, tem estimulado o desenvolvimento de uma série de ferramentas para gerar informações e novos conhecimentos a partir do processamento destes dados. Em razão disto, há um crescente interesse na aplicação de técnicas de mineração de dados em ambientes educacionais.

Este trabalho apresentou o levantamento do estado da arte de EDM na previsão de desempenho de alunos. Esta análise possibilitou conhecer melhor as técnicas de mineração utilizadas em trabalhos relevantes e atuais da área. Foi possível constatar que as árvores de decisão, apesar de ser um método de aprendizagem de máquina tradicional, é uma técnica bastante válida na mineração de dados educacionais.

Referências

Aviad, B. and Roy, G. (2011). Classification by clustering decision tree-like classifier based on adjusted clusters. *Expert Systems with Applications*, 38(7):8220–8228.

Baker, R. et al. (2010). Data mining for education. *International Encyclopedia of Education*, 7:112–118.

Castro, F., Vellido, A., Nebot, À., and Mugica, F. (2007). Applying data mining techniques to e-learning problems. In *Evolution of teaching and learning paradigms in intelligent environment*, pages 183–221. Springer.

Connolly, T. M. and Begg, C. E. (2005). *Database systems: a practical approach to design, implementation, and management*. Addison-Wesley Longman.

Han, J., Kamber, M., and Pei, J. (2006). *Data mining: concepts and techniques*. Morgan kaufmann.

Hanna, M. (2004). Data mining in the e-learning domain. *Campus-wide information systems*, 21(1):29–34.

Khatwani, S. and Arya, A. (2013). A novel framework for envisaging a learner's performance using decision trees and genetic algorithm. In *Computer Communication and Informatics (ICCCI), 2013 International Conference on*, pages 1–8. IEEE.

Koedinger, K., Cunningham, K., Skogsholm, A., and Leber, B. (2008). An open repository and analysis tools for fine-grained, longitudinal learner data. *Educational Data Mining*, 1:157–166.

Long, X. and Wu, Y. (2012). Application of decision tree in student achievement evaluation. In *Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on*, volume 2, pages 243–247. IEEE.

Merceron, A. and Yacef, K. (2005). Educational data mining: a case study. *Artificial Intelligence in education: supporting learning through Socially Informed Technology.–IOS Press*, pages 467–474.

Osmanbegović, E. and Suljić, M. (2012). Data mining approach for predicting student performance. *Economic Review*, 10(1).

Pellegrino, J. W., Chudowsky, N., Glaser, R., et al. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.

Quinlan, J. R. (1993). *C4.5: programs for machine learning*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.

Raghavan, N. S. (2005). Data mining in e-commerce: A survey. *Sadhana*, 30(2-3):275–289.

Romero, C., Espejo, P. G., Zafra, A., Romero, J. R., and Ventura, S. (2013). Web usage mining for predicting final marks of students that use moodle courses. *Computer Applications in Engineering Education*, 21(1):135–146.

Romero, C. and Ventura, S. (2010). Educational data mining: a review of the state of the art. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 40(6):601–618.

Romero, C., Ventura, S., and De Bra, P. (2004). Knowledge discovery with genetic programming for providing feedback to courseware authors. *User Modeling and User-Adapted Interaction*, 14(5):425–464.

Singh, S. and Kumar, V. (2012). Classification of student's data using data mining techniques for training & placement department in tecnical education. In *International Journal of Computer Science and Network (IJCSN)*, pages 121–126.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P., Zhou, Z.-H., Steinbach, M., Hand, D., and Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, 14(1):1–37.

Este documento foi organizado usando LaTeX 2ε