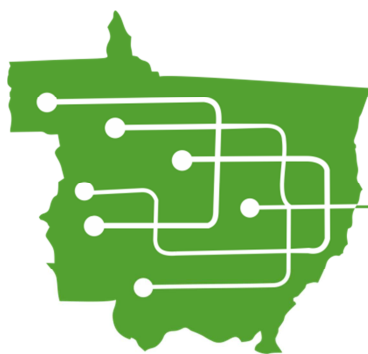




**Universidade Federal de Mato Grosso**  
**Campus Universitário do Araguaia**  
**Curso de Bacharelado em Ciência da Computação**  
**Sociedade Brasileira de Computação**



**Escola Regional de Informática**

SBC Mato Grosso

# **Anais da V Escola Regional de Informática de Mato Grosso - 2014**

Barra do Garças/ MT – 2014



## **V Escola Regional de Informática de Mato Grosso – ERI-MT 2014**

**Realização em 14 a 17 de outubro de 2014.**

ISSN: 2179-1953

### **Realização:**

- Universidade Federal de Mato Grosso - UFMT
- Curso de Bacharelado em Ciência da Computação – Campus Universitário do Araguaia

### **Promoção:**

- Sociedade Brasileira de Computação – SBC

### **Parceiros:**

- InforPoint Software
- DOS Informática
- PROCEV – UFMT
- Fundação Uniselva
- Autoposto Medalha
- Planos – Technology Division
- Certifiq

ISSN: 2179 – 1953

Copyright © 2013 da Sociedade Brasileira de Computação  
Todos os direitos reservados

Capa/Diagramação/Produção Editorial: Ivairton Monteiro Santos

**Editor**

Prof. Dr. Ivairton Monteiro Santos, Universidade Federal de Mato Grosso - UFMT

**Realização**

Universidade Federal de Mato Grosso - UFMT  
Curso de Bacharelado em Ciência da Computação

**Promoção**

Sociedade Brasileira de Computação – SBC

**Dados Internacionais de Catalogação na Fonte**

E74a Escola Regional de Informática (5. : 2014 : Barra do Garças, MT).  
Anais da V Escola Regional de Informática de Mato Grosso - 2014  
[recurso eletrônico] / Editor Ivairton Monteiro Santos. – Barra do Garças:  
Sociedade Brasileira de Computação: Universidade Federal de Mato  
Grosso. Curso de Ciência da Computação, 2014.  
1 CD-ROM ; 4 ¾ pol.

ISSN 2179-1953.

1. Informática. 2. Computação - Mato Grosso. I. Universidade Federal  
de Mato Grosso. Curso de Ciência da Computação. II. Sociedade  
Brasileira de Computação. III. Santos, Ivairton Monteiro (Ed.). IV. Título.

CDU 004(817.2)(063)

Ficha Catalográfica elaborada pelo Bibliotecário Carlos Henrique T. de Freitas – CRB-1 2.234

## Comissão organizadora

Prof. Me. Thiago Pereira da Silva (UFMT/Araguaia) (Coordenador geral da ERI-MT 2014)  
Prof. Me. Anthony Ferreira La Marca (UFMT/Araguaia)  
Profª. Meª. Carine Rodrigues (IFMT/Barra do Garças)  
Prof. Dr. Ivairton Monteiro Santos (UFMT/Araguaia)  
Profª. Drª. Livia Lopes Azevedo (UFMT/Araguaia)  
Prof. Me. Maxwell Carmo (UFMT/Araguaia)  
Prof. Me. Rafael Teixeira Sousa (UFMT/Araguaia)  
Prof. Me. Robson Silva Lopes (UFMT/Araguaia)

## Comitê técnico-científico

Prof. Dr. Ivairton M. Santos (UFMT/Araguaia) (Coordenador do Comitê)  
Prof. Dr. Alex Pinto (UFSC)  
Profª. Drª. Andreia Gentil Bonfante (UFMT/Cuiabá)  
Prof. Dr. Arlenes Silvino da Silva (UFMT/Cuiabá)  
Prof. Dr. Carlos Eduardo Cugnasca (USP/POLI)  
Prof. Dr. Carlos Eduardo da Silva (UFRN/ECT)  
Profª. Drª. Claudia Aparecida Martins (UFMT/Cuiabá)  
Prof. Dr. Clovis dos Santos Júnior (UFMT/CUR)  
Profª. Dr. Cristiano Maciel (UFMT/Cuiabá)  
Profª. Drª. Eunice Pereira dos Santos Nunes (UFMT/Cuiabá)  
Prof. Dr. Frederico Lopes (UFRN)  
Prof. Dr. Gleber Nelson Marques (UNEMAT/Alto Araguaia)  
Prof. Dr. Jó Ueyama (USP/ICMC)  
Prof. Dr. João Manuel Ribeiro da Silva Tavares (Universidade do Porto/Portugal)  
Prof. Dr. Josiel Maimone de Figueiredo (UFMT/Cuiabá)  
Prof. Dr. Linder Cândido da Silva (UFMT/Araguaia)  
Profª. Drª. Livia Lopes Azevedo (UFMT/Araguaia)  
Prof. Dr. Luis Rigo Jr. (UFES)  
Prof. Dr. Mauricio Pereira (UFMT/Cuiabá)  
Prof. Dr. Nelcilenio Virgílio de Souza Araújo (UFMT/Cuiabá)  
Prof. Dr. Norian Marranghello (UNESP/São José do Rio Preto)  
Prof. Dr. Peter Zeilhofer (UFMT/Cuiabá)  
Profª. Drª. Patricia Cristiane de Souza (UFMT/Cuiabá)  
Prof. Dr. Raul Wazlawick (UFSC)  
Prof. Dr. Roberto Benedito de Oliveira Pereira (UFMT/Cuiabá)  
Profª. Drª. Rogéria Cristiane Gratão de Souza (UNESP/São José do Rio Preto)  
Prof. Dr. Ronaldo Luis Alonso (UFMT/Araguaia)  
Prof. Dr. Saulo Roberto Sodrê dos Reis (UFMT/Eng. Elétrica)  
Prof. Dr. Sergio Carvalho (UFG)

## Revisores Adicionais

Prof. Me. Anthony Ferreira La Marca (UFMT/Araguaia)  
Profª. Meª. Carine Rodrigues (IFMT/Barra do Garças)  
Prof. Me. Rafael Teixeira Sousa (UFMT/Araguaia)  
Profª. Me. Soraia Silva Prietch (UFMT/Rondonópolis)  
Prof. Me. Thiago Pereira da Silva (UFMT/Araguaia)

## **Apresentação**

A Escola Regional de Informática de Mato Grosso (ERI-MT) é um evento promovido pela Sociedade Brasileira de Computação (SBC) e realizado anualmente, desde 2009, por diferentes Instituições de Ensino Superior do estado. No ano de 2014 o curso de Ciência da Computação do campus Araguaia da Universidade Federal de Mato Grosso teve o privilégio de realizar a V ERI-MT na cidade de Barra do Garças-MT, no período de 14 a 17 de outubro.

A quinta edição da ERI-MT teve o tema “A Computação Integrando Competências para o Desenvolvimento Regional”, e discutiu a formação profissional, as diversas tecnologias e como integrá-las para atender às demandas do estado por soluções tecnológicas. A V ERI-MT disseminou tecnologias e pesquisas científicas, promoveu o debate e o intercâmbio de conhecimento de temas relevantes da área de computação, que estão em evidência no país, além de fortalecer o intercâmbio entre as IES, as empresas e o governo do estado de Mato Grosso.

**Prof. Thiago Pereira da Silva**

*Coordenador da V Escola Regional de Informática de Mato Grosso*

## Sumário

### Artigos Completos

|  | Página |
|--|--------|
| Anthony Ferreira La marca and Marcus Vinicius Brito da Silva, Windelmarcos Borba, Marco Aurelio Silveira. <b>Interação Natural com Kinect para o Auxílio no Ensino da Disciplina de Estrutura de Dados.</b>                    | 9      |
| Rachel Almeida. <b>Marketing de relacionamento através de softwares CRM: estudos de caso no governo, na indústria e nos serviços.</b>  | 15     |
| Mattyws Grawe, Andreia Bonfante, Josiel Maimone Figueiredo, Claudia Martins, Raphael Gomes, Cristhian Ferreira, Phellippe Pinheiro. <b>Harmonização de nomes de depositantes de patentes: Uma abordagem com Apache Lucene.</b> | 21     |
| Maykon Perin dos Reis, Diego Brock, Paulo Garcia, Josefson Souza, Nilton Hideki Takagi, Josiel Maimone Figueiredo. <b>Indicadores em um PMO de Projetos de Pesquisa.</b>   | 27     |
| Pedro Clarindo da Silva Neto, Rafaely Sampaio, Cheiene Oliveira. <b>Análise de simulação de protocolo de roteamento geográfico para VANET.</b>   | 33     |
| Fabio Ricardo da Silva, Paulo Vinicius dos Santos Benedito, Raphael Gomes, Josiel Maimone Figueiredo. <b>Automatização dos Algoritmos para Estimativa do Balanço de Energia Superficial S-SEBI e SSEB.</b>                     | 39     |
| Alan Rodrigo Marcon, Nielsen Simões. <b>Identificação de Eventos de Áudio para Monitoramento de Aves no Pantanal Mato-grossense.</b>   | 45     |
| Phellippe Pinheiro, Cristhian Ferreira, Mattyws Grawe, Raphael Gomes, Josiel Maimone Figueiredo, Andreia Bonfante. <b>Comparação do desempenho entre banco de dados não relacional e relacional em análise de patentes.</b>    | 51     |
| Diego Neves, Igor da Costa, Adriano Roberto dos Santos, Fernando Tanaka, Soraia Prietch. <b>Proposta de Redesenho do Catálogo Online da Biblioteca da UFMT.</b>  | 57     |
| Alexandre Tomasi, Luciana Lee. <b>Comparação de Metaheurísticas para o Problema de Clusterização Automática.</b>   | 63     |
| Elias da Silva, Valdemar Neto. <b>MODEFS: Towards a lightweight Model-Driven Environment for Formal Software Engineering.</b>  | 69     |
| Eduardo Rolim, Claudio Monteiro, Fabiana Ferreira Cardoso. <b>Inception: Uma solução de abertura de portas dinâmicas para o protocolo SMIP.</b>  | 75     |
| Luan Lima, Danilo Londero, Mathias Mantelli, Lívia Lopes Azevedo. <b>RoboCar: protótipo de robô móvel autônomo.</b>  | 81     |
| Gracyeli S. Souza, Ivairton Monteiro Santos, Peter Zeilhofer. <b>Sistema para registro e manipulação de dados de precipitação de chuva do projeto TRMM.</b>  | 87     |
| Luy Lucas Santana, Allan Oliveira, Josiel Maimone Figueiredo, Thiago Ventura, Claudia Martins. <b>Aplicação para busca de similaridade em séries temporais.</b>  | 93     |
| Franco Magri Lupo, Ivairton Monteiro Santos. <b>Análise da capacidade de comunicação dos sensores Radiuino.</b>  | 99     |
| David Willians Beserra, Felipe Fernandes, Francielly Falcão da Silva, Wagner de Lima, Rubens Paula da Silva, Alexandre Borba, Jean Araujo. <b>A Computação em Nuvem e os Desafios de Infraestrutura para e-Science.</b>        | 105    |
| Fabiano Rocha. <b>Identificação dos fatores de maior impacto de aplicações Mapreduce.</b>  | 111    |
| Felipe Oliveira, Daniel Costa, Glaydson Vale, David Willians Beserra, Jean Araujo.   | 117    |

**Fragmentação de Memória em Ambientes Linux.**

|   |     |
|---|-----|
| David Willians Beserra, Anderson Cavalcanti, Luciano Souza, Rubens Paula da Silva, Jean Araujo, Sérgio Mendonça. <b>Simulador do Processo de Instalação do Rocks Clusters com uso de Técnicas de IHC.</b> | 123 |
| Rodrigo Silva, Nilton Hideki Takagi, Renato Junio Santos, Josiel Maimone Figueiredo. <b>Metodologia de Gerenciamento para Projetos inovadores com interação Público-Privado.</b>                          | 129 |
| Alicenaira Almeida, Evaldo Silva, Daves Martins. <b>Automação do Processo de Implantação de Software.</b>   | 135 |

**Resumos Expandidos**

|   | Página |
|---|--------|
| Cesar Guarienti, Nielsen Simões. <b>A Implementação de monitoramento centralizado em servidores do Instituto de Computação da UFMT.</b>   | 142    |
| Bruna Carlota, Fernando Da Cruz, Andreia Bonfante. <b>A utilização de dicionário de sinônimos na análise automática de patentes.</b>  | 145    |
| Allan Alencar, Rafaely Lino, Andreia Bonfante. <b>A utilização de medida de similaridade na classificação automática de documentos de patentes.</b>   | 148    |
| Aline Souza, Rennan de Figueiredo, Lucas Sartorello. <b>Legado digital: novos paradigmas culturais.</b>   | 151    |
| Eduardo Obara, Devair Junior, Leonardo Souza, Jhony Souza, Eliel Lima. <b>Automação do cálculo de velocidade média e estudo de conceitos de física térmica e mecânica com um barco a vapor.</b>   | 154    |
| Gustavo Pereira da Silva Guedes, Kelven Douglas da Silva, Aline P. Domingos de Souza. <b>Automação Residencial: Buscando melhorias na acessibilidade de baixo custo.</b>                          | 157    |
| Aline Souza, Marlon Bordin. <b>Computação Gráfica na Educação Pública Nacional.</b>   | 160    |
| Eduardo Lemos, Eunice Nunes, Cristiano Maciel. <b>Uma Revisão de Literatura sobre Recursos de Interação Tridimensional para embasar o desenvolvimento de Brinquedo Terapêutico Digital.</b>       | 163    |
| Mathias Mantelli, Rafael Teixeira Sousa. <b>Avaliação de métodos de combinação de classificadores em um sistema de auxílio ao diagnóstico de pneumonia infantil.</b>                              | 166    |
| Ketelem Lemos de Campos, Bruno Honorato de Oliveira, André Luiz Miranda, Cleder Almeida, Cristiano Maciel. <b>e-Con: Proposta de software Web para o Procon.</b>                                  | 169    |
| Thassiany Pires, Vinicius Silva, Fabiano Rocha. <b>Automação Residencial por comando de voz em Delphi 7.</b>  | 172    |
| Marcos Danillo Lau, João Pedro Pereira, Leandro Ferrarezzi, Ricardo Oliveira, Fabiano Rocha. <b>Motores de Passo: estudo e aplicabilidade em uma maquete interativa controlada em tempo real.</b> | 175    |
| Luan Lima, Walas Moraes, Maxwell Carmo. <b>Uma Arquitetura para Gestão da Largura de Banda em Redes Domésticas.</b>   | 178    |
| Lucas Borges, Claudia Leite, Fabiano Rocha, Wellington Cruz, Gabriel Oliveira, Vinicius Ferri. <b>Inclusão: Glossário Interativo de Libras.</b>   | 181    |
| Lyncon Lima, Roan Siviero Dotto, Leticia Cristina Amorim de Assis, Abraão Lincon Ramalho, Bruno Silva Macário, Juliana Antunes. <b>Proposta da utilização do método</b>                           | 184    |

|   |     |
|---|-----|
| <b>Goal/Question/Metric para avaliação de acessibilidade em uma Biblioteca Virtual.</b>   |     |
| Walas Moraes, Luan Lima, Maxwell Carmo. <b>Acesso Compartilhado à Internet através de Redes WiFi.</b>   | 187 |
| Carine da Costa, Lorena Leal Oliveira, Thatielle M. Pereira, Beatriz de Paula Moraes. <b>Usabilidade nos Portais dos Institutos Federais do Centro-Oeste.</b>         | 190 |
| Pedro Francisco de Sousa Neto, Marcelo Quinta, Mariângela Estelita Barros. <b>ELiS Mobile: trazendo a comunicação escrita dos surdos para os dispositivos móveis.</b> | 193 |

**Caderno de**  
**Artigos Completos**



## Interação Natural com Kinect para o Auxílio no Ensino da Disciplina de Estrutura de Dados

**Anthony F. La Marca, Marcus Vinícius B. Da Silva, Windelmarcos F. De Borba, Marco Aurélio M. Silveira**

Instituto Ciências Exatas e da Terra (ICET) – Universidade Federal de Mato Grosso (UFMT) - Barra do Garças – MT.

(anthonyferreiralamarca, marcusbrito17, windelmarcos)@gmail.com  
marco\_aurelio\_ro@hotmail.com

***Abstract.** Since the Human-Computer Interactions studies started growing, different tools were created in order to provide users a greater ease of interactions with machines. Those tools have great potential due to the new associated technologies. All things considered, studies in the academic environment become viable. This work discusses the creation of an application using the Kinect device to assist in data structure teaching, specially since this discipline has a great number of dropout students at the Federal University of Mato Grosso, region of Araguaia.*

***Resumo.** Com o surgimento de estudos na área de Interação Humano-Computador, foram criadas diversas ferramentas que proporcionavam aos usuários maior facilidade de interação com as máquinas. Tais ferramentas possuem grande potencial, devido as novas tecnologias associadas, tornando-as viáveis aos estudos no meio acadêmico. Desta forma, o trabalho aborda a criação de uma aplicação utilizando o dispositivo Kinect para auxiliar no ensino de estrutura de dados, visto que, a disciplina apresenta um número grande de reprovações e desistências de alunos na Universidade Federal do Mato Grosso, campus Araguaia.*

### 1. Introdução

O aprendizado não é apenas em uma sala de aula com professores e livros, está em todos os momentos em que o ser humano passa por novas experiências. Em muitos casos, os métodos tradicionais não são tão eficazes, pois cada aluno possui suas próprias características, interesses, capacidades e necessidades de aprendizagem. Desta forma, faz-se necessário desenvolver um meio onde alcance o interesse de todos, para que assim, tenham um melhor aproveitamento do ensino.

Um dos grandes diferenciais observados atualmente no ensino é o de uso de tecnologias computacionais, na qual consegue-se maior atenção dos alunos por oferecerem recursos que aumentem o realismo para/com o ensino. Neste contexto, rastreadores de movimentos “casados” com aplicações de Visualização da Informação, surgiram para trazer inovações ao mundo de pesquisa e de ensino.

Entre os rastreadores disponíveis, o Kinect lançado em 2010, proporcionava aos usuários um comportamento livre, sem a necessidade de controle, apenas utilizando o próprio movimento do corpo, diferentemente de outros lançamentos da época. O que este rastreador de movimento proporciona aos usuários, chamamos de Interação

Natural, sendo um sistema capaz de entender ações naturais do ser humano e repassa-lás em tempo real para o ambiente tridimensional, modificando-o. (Valli, 2005).

Após o lançamento do Kinect, surgiram APIs (*Application Programming Interface*) não oficiais que auxiliavam no desenvolvimento de aplicações de Interação Natural com o dispositivo. Com o grande sucesso destas aplicações, a Microsoft lançou uma versão oficial do Kinect para o sistema operacional *Windows*, além de disponibilizar um SDK (*Software Development Kit*) gratuito, emergindo estudos com o dispositivo na área de pesquisa e ensino.

Atualmente observa-se grande dificuldade no ensino da disciplina de estrutura de dados, visto que, o número de desistências e de reprovações de alunos é grande. Este fato é observado na Universidade Federal do Mato Grosso, campus Araguaia (Siga, 2014), assim, para minimizar o problema, o trabalho visa em desenvolver uma aplicação de Interação Natural com o dispositivo Kinect e criar uma ferramenta para auxiliar no ensino da disciplina proposta. O foco do trabalho é ensinar a parte teórica de estrutura de dados, assim técnicas e fundamentos de linguagem de programação indifere na utilização do sistema, ficando a cargo do professor da disciplina tal responsabilidade.

Inicialmente apenas as estruturas pilha e fila foram implementadas, as demais estruturas serão incorporadas ao trabalho futuramente.

## **2. Interação Humano-Computador (IHC)**

Interação Humano-Computador surgiu por volta de 1980, junto com a criação dos computadores pessoais. É uma área multidisciplinar composta por psicologia, ciência da computação, linguísticas, artes, dentre outras, com o objetivo de criar uma melhor forma de comunicação entre o ser humano e computadores (Valli, 2005).

É comum que o ser humano utilize o auxílio da linguagem gestual para se comunicar, por ser uma forma natural e simples. Com isso será utilizado na aplicação a NUI (Natural User Interface) por meio do Kinect, que é um dos mais novos modelos para criar interfaces com interação natural, onde procura-se criar sistemas que entendem ações que são naturais aos usuários, como gestos, expressões e movimentos (Valli, 2005).

Gerar uma aplicação de Interação Natural se tornou de fácil acesso, devido aos diversos dispositivos de baixo custo existentes no mercado e a disponibilização de ferramentas para desenvolvedores dedicadas para esses dispositivos.

## **3. Dispositivo de Rastreamento - Kinect**

Os dispositivos de rastreamento são utilizados para acompanhar a posição do corpo e os movimentos do usuário, assim como a posição de outros objetos por ele utilizados. Após a captação destes movimentos, eles são repassados ao mundo tridimensional modificando-o em tempo real, além do *feedback* para o usuário.

Um destes dispositivos é o Kinect, lançado em 2010 pela Microsoft com o objetivo de mudar a forma de interação das pessoas com os jogos, sendo uma tecnologia onde as pessoas não teriam necessidade de utilizar controle físicos para jogar, mas sim, o próprio movimento do corpo.

O Kinect tornou-se uma ferramenta de grande potencial para aplicações com IN (Interação Natural) em diversas áreas. Dentre os âmbitos que o Kinect está sendo inserido, além do de jogos, há pesquisas na área de medicina, de engenharia, de segurança, e principalmente na de ensino, que é a abordagem deste trabalho (Medeiros, 2012).

#### **4. Estrutura de Dados**

O mundo computacional, desde sua criação durante a segunda guerra mundial, onde surgiram os primeiros computadores conhecidos atualmente, sofre constantes evoluções criando novas perspectivas. Como consequência dessas evoluções, a quantidade de dados a serem manipulados e armazenados teve um aumento extraordinário, tornando a eficiência de como controlá-los um fator de grande importância. Para resolver esse problema, começaram a criar novas técnicas para melhorar o aproveitamento de memória e de processamento para o controle desses dados. Essas técnicas foram chamadas de estrutura de dados, que segundo Laureano (2012), nada mais é do que formas de armazenamento utilizadas no mundo real adaptadas para o mundo computacional com a finalidade de otimizar a manipulação e o tratamento de dados eletronicamente.

#### **5. Descrição do Trabalho**

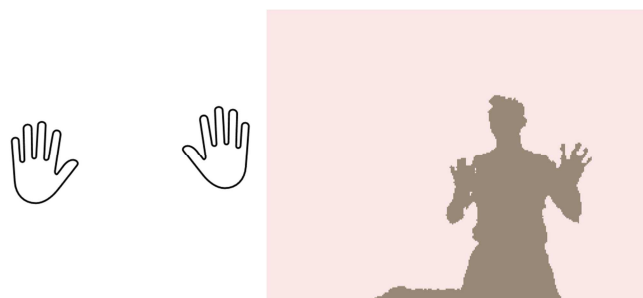
Com a finalidade de suprir os objetivos apresentados na introdução, a ferramenta propõe o desenvolvimento de uma aplicação capaz de oferecer uma alternativa na metodologia de ensino do professor para os alunos no ensino de estrutura de dados, visto que, o sistema utiliza componentes gráficos e um dispositivo de rastreamento para a interação dos usuários para/com o sistema. Desta forma, disponibilizar um complemento tecnológico para o ensino de estrutura de dados, estimulando os envolvidos através de imagens computadorizadas e equipamentos não convencionais, tenta-se minimizar o número de reprovações e desistência da disciplina.

O trabalho utiliza a API da Microsoft denominada Kinect SDK (Microsoft, 2014) para realizar toda a interação natural do sistema, desde na manipulação da aplicação até mecanismos de inserção e de remoção nas estruturas propostas. A documentação de toda a API está disponível em (Cardoso, 2013).

As principais tarefas do sistema é a inserção e a remoção de elementos nas estruturas pilha e fila, além de representações gráficas e comandos via interação natural.

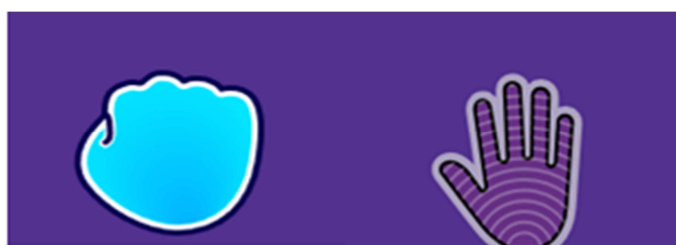
Para a interação natural com a aplicação utilizando o Kinect SDK é necessário aplicar o fluxo de interação, que fornece todas as opções de objetos para a interface gráfica específicas para a manipulação com o Kinect.

Dentro da área do KinectRegion (uma classe da API), onde os objetos estão contidos, é capturado automaticamente o posicionamento de uma das mãos do usuário à frente do Kinect. Já nessa aplicação foram feitas alterações na DLL (“*Dynamic-Link Library*” ou “Biblioteca de Vínculo Dinâmico”) responsável pelo fluxo de interação para que seja possível o rastreamento das duas mãos do usuário simultaneamente, conforme ilustra a Figura 1.



**Figura 1 – Interação com Aplicação Utilizando as Mãos**

A interação com um botão, ou seja, pressioná-lo é feita por meio do movimento de empurrar a mão em direção ao sensor Kinect, dando a impressão de estar realmente apertando um objeto. Na KinectScrollViewer, ou barra de rolagem do Kinect, é manipulada por meio do fechamento da mão. Após fechar a mão é possível arrastar a barra na direção desejada, conforme ilustra a Figura 2.



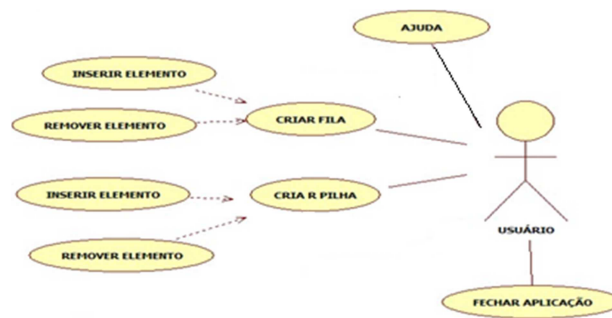
**Figura 2 – Arrastando e Pressionando**

O trabalho utilizou o algoritmo especializado chamado keyframes. O método keyframe utiliza quadros-chaves, onde é definido valores a certos parâmetros que variam com o tempo em cada movimento, como posição ou cor. Por meio desses valores é calculada a evolução dos parâmetros em cada quadro de imagem em um determinado tempo (Gomide et al 2009). Este limite entre os parâmetros do quadro inicial até o quadro no tempo final é utilizado para identificar a velocidade necessária para que o movimento seja aceito como um comando na aplicação (Cardoso 2013).

Na aplicação os movimentos utilizados para executar certos comandos são chamados de rastreadores. Esses rastreadores definidos na aplicação são os utilizados para inserção e remoção de valores na pilha e na fila, além de serem responsáveis por retornar à tela principal.

A classe do rastreador responsável pelo movimento de inserção nas estruturas é feita por um movimento de aceno com o braço direito. Esse movimento é composto por uma lista de quadros com três poses, onde é necessário que inicialmente a mão direita do usuário esteja em uma posição maior que a posição do cotovelo direito em um eixo horizontal. Na pose seguinte, a mão e o cotovelo do usuário devem estar em uma mesma posição no eixo horizontal e a mão em uma posição superior em relação ao eixo vertical, e como última pose, a mão direita deve estar em uma posição inferior com relação ao eixo horizontal. Já o movimento de remoção é similar, mas deve ser executado utilizando o braço esquerdo. Com relação ao movimento para retornar ao menu principal, basta posicionar os braços à frente e baixa-los.

A Figura 3 ilustra o Diagrama de Caso de Uso do sistema, descrevendo suas funcionalidades, ou seja, apresentando a sequência de eventos do ator que utiliza o sistema para montar as estruturas e analisar as operações sobre elas.



**Figura 3 – Diagrama de Caso de Uso**

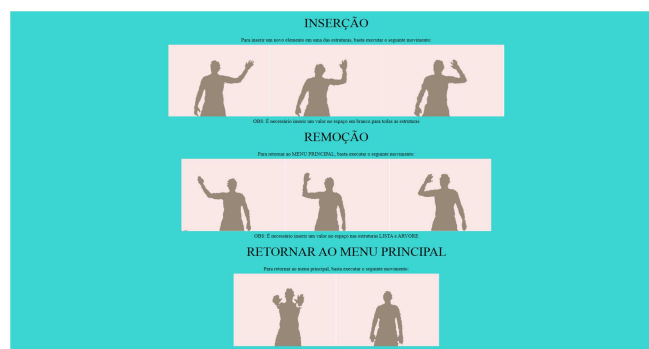
## 6. Resultados

Como resultados finais no desenvolvimento da aplicação, foi obtida uma interface simples e de fácil manuseio para o usuário, por meio de interações naturais. Efetuando movimentos diante do Kinect é possível utilizar toda a funcionalidade da aplicação, dispensando a utilização de dispositivos convencionais, como mouse e teclado. A Figura 4 ilustra a interface inicial do sistema



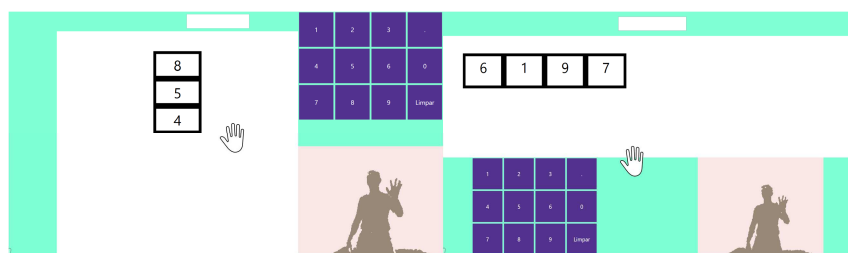
**Figura 4 – Tela do Menu Principal**

Na interface gráfica do *menu principal* é disponibilizado botões para selecionar as opções de temas disponíveis relacionados a disciplina de estrutura de dados. Dentre os temas à serem escolhidos estão a pilha e a fila, além de estar disponível a opção de ajuda (Figura 5), que identifica todos os comandos possíveis da aplicação e como executá-los.



**Figura 5 – Tela de Ajuda**

Após a escolha de uma das estruturas é mostrada a sua respectiva interface. A Figura 6 apresenta as interfaces das estruturas pilha e fila.



**Figura 6 – Estrutura Pilha e Fila**

## 7. Conclusão

O escopo do trabalho foi utilizar técnicas que garantissem fácil acesso à ferramenta pelas comunidades interessadas, disponibilizando um ambiente integrado na qual a partir dele, seja capaz de interagir com grande parte das estruturas utilizadas na disciplina proposta.

Assim, concluí-se que a aplicação auxilia no ensino da disciplina de estrutura de dados, pois proporciona maior interação dos usuários com o sistema, por utilizar recursos tecnológicos na representação de conceitos envolvidos na disciplina proposta, aumentando o interesse e o envolvimento dos alunos na aprendizagem. Além de estimular e melhorar o aproveitamento dos alunos, tornou-se uma aplicação que possibilita uma alternativa na metodologia de ensino para os professores da disciplina, por conter mecanismos suficientes para a execução de aulas voltadas aos assuntos abordados pela aplicação.

Como trabalhos futuros, novas estruturas podem ser incorporadas ao trabalho, devido a sua programação modular, além da substituição do dispositivo Kinect pelo novo lançamento da Microsoft (2014), chamado Kinect One, que proporciona maior sensibilidade aos movimentos.

## 8. Referências

- Cardoso, Gabriel, S. Microsoft Kinect: Crie aplicações interativas com o microsoft kinect. Casa do Código. São Paulo, 2013.
- Gomide, João V. b; Pacheco, Daniel; Araújo, Arnaldo de a. Captura de movimento e animação de personagens em jogos. Universidade Fumec e Universidade Federal de Minas Gerais. Rio de Janeiro – RJ, 2009.
- Laureano, Marcos. Estrutura de dados com algoritmos e c. Curitiba: Brasport livros e multimídia ltda, 2012.
- Microsoft. Kinect for windows sdk. disponível em: < <http://msdn.microsoft.com/en-us/library/hh855347.aspx> >. Acesso em: julho de 2014.
- Medeiros, Anna Soares. Interação natural baseada em gestos como interface de controle para modelos tridimensionais. Universidade Federal da Paraíba. João Pessoa, 2012.
- Siga. Sistema de Informações de Gestão Acadêmica. Disponível em: <https://sia.ufmt.br/www-siga/htmldelphi/login/LoginUnico/login.htm>. Acesso em: agosto de 2014.
- Valli, A. Notes on natural interaction. 2005.

# Marketing de relacionamento através de software CRM: estudos de caso no governo, na indústria e nos serviços

Rachel de Camargo Serpa de Almeida<sup>1</sup>

<sup>1</sup>FASB - Faculdade do Sul da Bahia  
Teixeira de Freitas - BA - Brasil.

rachel.almeida@ffassis.edu.br

**Abstract.** *This paper aims to present the fundamental concepts on which are based the CRM software and contains case studies described in recent Brazilian scientific papers that analyze the application of this type of system in government, industrial and service organizations, culminating with the example application of CRM in a private institution of higher education.*

**Resumo.** *O presente trabalho tem como objetivo apresentar os conceitos fundamentais sobre os quais se fundamentam os software CRM, apresentando ainda estudos de caso descritos em trabalhos científicos brasileiros recentes que analisam a aplicação desse tipo de sistema em organizações governamentais, industriais e de serviços, culminando com o exemplo de aplicação de solução CRM em uma instituição de ensino superior privada.*

## 1. Introdução

A evolução do relacionamento entre empresas e clientes acompanha o desenvolvimento da economia mundial. Globalização, revoluções tecnológicas e aumento da competitividade são alguns dos fatores que levam ao Marketing o desafio de atrair, reter e fidelizar uma base de clientes cada vez mais exigente e exposta a múltiplos produtos e serviços. Tal realidade põe à prova sistemas como *push/pull*, que tentam adaptar o cliente à oferta ou produto, passando da máxima "qualquer cor, desde que seja preto" para "diga que cor prefere" (MCKENNA, 1992). Nesse sentido, surge na década de 1990 o conceito de CRM, sigla indicativa de *Customer Relationship Management*, que pode ser traduzida como gestão do relacionamento com o consumidor.

A questão é que o conceito de CRM, na prática, está intimamente ligado à tecnologia da informação, sendo tema recorrente em publicações científicas direcionadas à área. Este artigo apresenta estudo bibliográfico sobre aplicações do CRM no governo, na indústria e nos serviços bancários, culminando com o caso real de implementação em uma instituição de ensino superior privada do extremo sul baiano.

## 2. Referencial Teórico

A gestão do relacionamento com clientes tem por objetivo administrar as informações sobre os consumidores para, segundo Kotler (2012), desenvolver significados e ofertas específicas de produtos e serviços centradas no consumidor, oferecendo incentivos motivacionais para estimular a repetição da compra e aumentar a fidelização. O contexto histórico em que o CRM se desenvolve se caracteriza por duas tendências centrais: o crescimento da competitividade global e o aprimoramento na captura e administração de informações (*database*) promovido pela evolução das ferramentas tecnológicas.

Através da interatividade proposta pelo CRM é possível descobrir e tratar informações valiosas a respeito de clientes, produtos e alternativas de adaptação a novas frentes de mercado. A consequência direta disso é o aumento das oportunidades de *cross-selling* (venda cruzada), embora múltiplas outras utilidades possam ser extraídas desse tipo de sistema, que representa uma transformação de antigos conceitos em ideias práticas, materializadas na forma de novas tecnologias capazes de aproximar o pensamento estratégico da efetiva ação tática comercial.

## 3. Aplicações do CRM

### 3.1. Renovação democrática: o CRM no governo britânico

A pesquisa promovida por Batista (2006) avaliou a efetividade de um sistema CRM em esfera municipal na Inglaterra, implantado no sentido de otimizar a capacidade de resposta e ampliar o desempenho geral do governo local. O autor observa que novas e múltiplas expectativas geradas pelos cidadãos frente ao governo puderam ser tratadas, na retaguarda, através de sistemas capazes de entender tais necessidades e promover, com isso, maior comunicação, transparência e agilidade na tomada de decisões. No caso governamental o termo “cliente” é amplo, referindo-se aos cidadãos, empresas comerciais diversas, associações comunitárias, entidades militares e religiosas, sindicatos e outros caracterizados como *stakeholders* dos serviços públicos locais.

Embora a adoção de sistemas CRM por parte dos governos ainda seja incipiente, o artigo sugere que tais soluções são capazes de prover maior segurança aos administradores públicos. A partir do entendimento sobre as reais estratégias de aproximação com a população atendida, os maiores desafios envolvem: o alinhamento da estrutura operacional e da cultura governamental ao conceito de CRM, que precisa ser adaptado das organizações empresariais tradicionais; a dificuldade de compartilhamento das informações dentro do governo; e as estratégias de integração entre diferentes canais de contato / interação e entre sistemas legados e novas tecnologias.

### 3.2 Impactos financeiros: o CRM em uma indústria farmacêutica

O trabalho de Kazato *et all* (2008) trata o CRM sob o ponto de vista do efetivo retorno financeiro auferido por uma empresa que implementa esse tipo de solução. A pesquisa avalia uma unidade de negócios de empresa industrial do segmento farmacêutico. Para os autores, existe uma dificuldade importante em evidenciar tal ganho, que pode ser apontado como um dos principais motivos de fracasso do processo. No estudo de caso foi avaliada uma empresa que já dispunha de sistema CRM anterior, que havia sido recentemente substituído por outro, mais moderno.

Como conclusão os autores verificaram a existência de vantagens associadas a ambas as ferramentas. A motivação para o investimento em novas tecnologias CRM se relacionou a quatro fatores básicos: interface, uma vez que o novo modelo apresentava-se mais familiar aos usuários; compartilhamento de informações, capaz de trazer para o campo explícito o conhecimento até então restrito à dimensão tácita; simplificação das atividades cotidianas; e flexibilidade trazida por múltiplas possibilidades de *layout* configuráveis pelo usuário de acordo com seus interesses e preferências.

A pesquisa conclui que a grande quantidade de informações armazenada dificultou o conhecimento aprofundado sobre todas as possibilidades que o sistema apresenta, e que o declínio da produtividade em uma organização pode estar associado, justamente, ao excesso de informações. Já quanto ao impacto no volume de vendas de produtos, os autores observaram impactos positivos produzido pelo CRM nos indicadores financeiros, que se elevaram em no mínimo 3,1%. Enfim, Kazato *et all* atribuem grande parte do sucesso do CRM ao fato de que os funcionários consideravam que os conceitos que envolviam o software estavam plena e amplamente esclarecidos antes de sua efetiva implantação.

### 3.3 CRM em serviços bancários: impactos na iniciação, manutenção e finalização de relacionamentos comerciais

O artigo de Melo *et all* (2009) analisam a aplicação de soluções CRM no setor de serviços bancários, compreendendo como tais sistemas favorecem a avaliação de clientes atuais e potenciais através de um maior volume de informações sobre os mesmos e suas interações. A pesquisa, conduzida em uma instituição bancária brasileira, tem foco nos impactos do CRM em três fases do relacionamento empresa – cliente: iniciação; manutenção; e finalização.

Os autores constataram que, para o consumidor, cada ponto de contato com o banco representa uma relação com a empresa, e a memória a respeito dessas relações pode vir a definir a continuidade ou não de um padrão de consumo. Os sistemas CRM buscam, através da coleta e análise de dados, possibilitar o estabelecimento de relações de confiança tipo ganha-ganha entre a base de clientes e a empresa (GORDON, 1999). A partir do CRM é possível gerenciar todo o ciclo de vida da relação empresa-cliente, e pode tratar até mesmo as relações com aqueles que pretendem deixar ou já deixaram a base.

Segundo pesquisas do Gartner Group (2003), ainda que as vantagens do CRM sejam inquestionáveis, os resultados organizacionais são insatisfatórios em 70% das implementações. Mais especificamente no setor de serviços bancários, alguns itens preocupantes devem ser considerados: multiplicidade de clientes; extensa carteira de produtos; e a quantidade elevada de canais de atendimento.

A pesquisa qualitativa levantou os principais efeitos do uso do CRM em cada fase do ciclo de vida do relacionamento entre banco e clientes. As conclusões apontam para uma efetiva melhoria de desempenho na relação com clientes na fase de manutenção, representada principalmente pelas oportunidades de *cross-selling* (venda cruzada) geradas. Já durante a inicialização o sistema foi ineficiente, uma vez que trata apenas de interações anteriores, sendo incapaz de identificar clientes em potencial caso esses não pertençam à base atual. Na finalização os autores constataram que, embora o sistema seja capaz de identificar clientes inativos, o mesmo não é utilizado para isso. Por fim, Melo *et all* (2009) apontam para a manutenção de segmentação de clientes por critérios tradicionais de renda e investimento, mesmo após a implantação do CRM, não extraindo da solução sua potencialidade máxima.

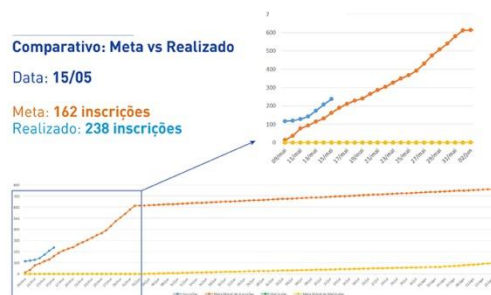
### **3.4 Estudo de caso: solução CRM em uma instituição privada de ensino superior**

A partir dos resultados encontrados em trabalhos científicos aqui discutidos é possível pressupor que, de fato, há ganho relativo na implementação do CRM nos mais diversos contextos. Os principais desafios parecem apontar para o comprometimento das equipes envolvidas, extração e tratamento de informações relevantes e adaptação aos interesses da organização e seus usuários.

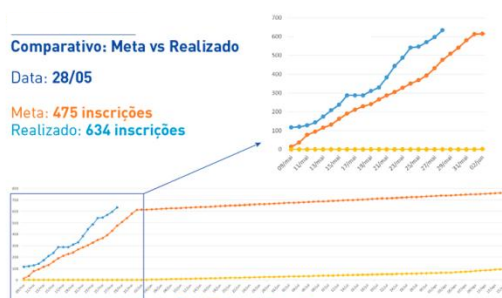
A aplicabilidade da solução foi testada em pesquisa de campo, de caráter quantitativo, realizada em uma instituição de ensino superior privada, localizada no extremo sul da Bahia, entre os meses de maio e julho de 2014, período de implantação do sistema CRM. A motivação principal da mesma nesse investimento foi desenvolver uma estratégia de captação, retenção e fidelização de clientes através da "diferenciação em massa", ou seja, o entendimento do cliente em sua individualidade, com suas preferências, padrões de comunicação e escolhas devidamente identificados e tratados. O primeiro desafio do sistema foi no suporte ao vestibular de inverno da instituição, tradicionalmente considerado inviável economicamente. Avaliando a relação entre investimento em mídia e operacionalização do vestibular e efetiva captação de estudantes para os cursos de graduação, o vestibular de inverno já havia sido abandonado há 3 anos.

Na ação pesquisada, o sistema foi alimentado com dados de vestibulares anteriores, perfis de ingresso e origem da base de dados de alunos que realizaram provas de vestibular nos últimos 3 anos. A partir da identificação de perfis específicos de clientes potenciais que buscavam a educação superior fora do período de vestibular tradicional, ações totalmente direcionadas a esse público foram executadas por ocasião do vestibular de inverno de 2014. Cabe ressaltar que, em paralelo ao CRM, houve investimento maciço em capacitação de funcionários para a conscientização sobre a importância da nova estratégia.

Os resultados podem ser demonstrados pelos gráficos abaixo:



**Gráfico 1 - Inscrições para o vestibular de inverno no início da campanha. Fonte: dados da pesquisa de campo.**



**Gráfico 2 - Inscrições para o vestibular de inverno após conscientização e treinamento dos funcionários. Fonte: dados da pesquisa de campo.**



**Gráfico 3 - Resultado geral de inscrições para o vestibular e início do período de matrículas. Fonte: dados da pesquisa de campo.**

A definição das metas, tanto em inscrições para o vestibular quanto em matrículas efetivadas, partiu de um patamar considerado agressivo, tendo em vista as condições populacionais, econômicas e culturais em que a região do extremo sul baiano está inserida. O último resultado parcial, apurado em 30/07/2014, dava conta de um ganho real de 59% na efetivação de matrículas frente à meta estipulada, e de 53% quanto a adesão ao vestibular de inverno. Em paralelo à admissão de novos alunos, o fato de um potencial cliente cadastrar-se para a realização do vestibular já oferece para a instituição, via CRM, possibilidades múltiplas de contato para a oferta de serviços educacionais *a posteriori*.

A partir dos dados coletados é possível verificar que, a partir de uma solução CRM, é possível de fato auferir resultados positivos relevantes desde a sua implementação. No caso da instituição de ensino superior pesquisada, foi notório ainda perceber que, alinhado a aplicação da ferramenta, constituiu-se importante diferencial o

estabelecimento de uma estratégia alinhada à cultura organizacional, contando com a adesão de todos os funcionários ao processo.

#### **4. Considerações Finais**

O conceito de CRM constitui-se em uma nova terminologia para antigos conceitos, alinhada a um grande suporte tecnológico. Mais que um mero produto, o CRM deve ser entendido como parte da estratégia organizacional, tornando-a mais próxima de seus clientes atuais e potenciais.

As pesquisas referenciadas nesse artigo demonstram que, em diversos segmentos, as organizações que extraem melhores resultados do CRM são aquelas que encaram a solução a partir de uma perspectiva estratégica, integrada e alinhada à cultura, em detrimento de uma visão tecnicista de mera aplicação de software. O grande desafio que se apresenta é: como explorar devidamente as múltiplas possibilidades geradas pelo CRM nas relações empresa-cliente, em suas múltiplas facetas durante o ciclo de vida? Independente do setor econômico, aparentemente há uma convergência em torno do assunto, e as organizações buscam planejar e sistematizar ações que usufruam de todo o potencial tecnológico citado, reconhecendo suas limitações, avaliando resultados e identificando oportunidades sempre que possível.

#### **Referências**

- Batista, L. (2006) "Aplicação de sistemas CRM no governo: análise de uma instituição pública britânica". In: Anais do 3º CONTECSI, São Paulo.
- Gordon, I. (1999), Marketing de Relacionamento: estratégias, técnicas e tecnologia para conquistar clientes e mantê-los para sempre. Ed. Futura, São Paulo.
- Kazato, P., Moretto, L. e Schmitt, V. (2008) "Contribuição de um Sistema de Gerenciamento de Relacionamentos com Clientes (CRM) nos Resultados Financeiros: estudo de caso da indústria farmacêutica z". In: Anais do 5º CONTECSI, São Paulo.
- Mckenna, R. (1992), Marketing de Relacionamento: estratégias bem-sucedidas para a era do cliente. Ed. Campus 8ª ed., Rio de Janeiro.
- Melo, F.A.F., Freitas, A.A.F. e Melo, C.F (2009) "Análise das Aplicações de Sistemas CRM na Iniciação, Manutenção e Finalização de Relacionamentos Comerciais em Provedores de Serviços Bancários". In: Anais do 6º CONTECSI, São Paulo.
- O'Brien, J.A. e Marakas, G.M (2013), Administração de Sistemas de Informação. Ed. Bookman, McGraw-Hill 15ªed. Porto Alegre.
- Swift, R. (2001). CRM, Customer Relationship Management: o revolucionário marketing de relacionamentos com os clientes. Ed. Elsevier, Rio de Janeiro.

# Harmonização de nomes de depositantes de patentes: Uma abordagem com Apache Lucene™

Mattyws Ferreira Grawe<sup>1</sup>, Andreia Gentil Bonfante<sup>1</sup>,  
Josiel Maimone de Figueiredo<sup>1</sup>, Claudia Aparecida Martins<sup>1</sup>,  
Cristhian M. Ferreira<sup>1</sup>, Phellippe Pinheiro<sup>1</sup>, Raphael de Souza Rosa Gomes<sup>1</sup>

<sup>1</sup> Instituto de Computação - Universidade Federal de Mato Grosso (UFMT)

{mattyws, cmferreira.352, phellippe.end}@gmail.com

{andreia.bonfante, josiel, claudia, raphael}@ic.ufmt.br

**Abstract.** *This paper presents an implementation to a patentee name harmonization module, using text mining techniques, focusing on treatment of common name descriptors, generation of acronyms and name condensing. This module will be part of the Periscope, an open source patent analysis tool.*

**Resumo.** *Este artigo apresenta resultados iniciais da re-implementação de um módulo de harmonização de nomes de depositantes de patentes, utilizando técnicas de processamento de texto, dentre as quais a identificação e retirada de descritores comuns e a utilização de acrônimos. Essa funcionalidade faz parte do Periscope, uma ferramenta baseada em software livre que auxilia na análise de patentes.*

## 1. Introdução

Do ponto de vista de inovação tecnológica, os documentos de patentes estão entre as melhores opções de fontes de consulta, pelo fato de possuírem maior detalhe técnico, organizacional e de serem de fácil acessibilidade. Análises desses documentos são necessárias para se manter a competitividade do mercado e também fomentar novas pesquisas [Abbas et al. 2014].

O primeiro quesito para se obter estatísticas e mapeamentos de qualidade é se ter uma base de dados confiável. A falta de uniformização de nomes de depositantes e inventores traz grandes problemas nessas avaliações. Por se tratar de informações vindas de várias fontes, erros tipográficos, acrônimos e variações de nomes podem ocorrer. Por exemplo, é comum se encontrar nas bases casos como esses: dois nomes, “BAYER ANIMAL HEALTH GMBH” e “BAYER ANIMLA HEALTH GMBH”, que indicam a mesma empresa, porém com as letras *A* e *L* trocadas em *ANIMAL*. Em um processo de coleta automática de dados, esses dois nomes podem ser considerados como de entidades distintas, resultando em análises imprecisas.

Assim, torna-se necessário que haja uma harmonização desses nomes. Isso é feito com a utilização de técnicas de processamento automático de texto [do Prado and Ferneda 2008], capazes de identificar essas divergências.

Neste artigo, são apresentados alguns resultados iniciais obtidos na implementação da segunda versão de um módulo responsável por produzir essa

harmonização. O módulo faz parte de uma aplicação, chamada Periscope, uma ferramenta de análise de patentes, que está sendo desenvolvida pelos autores <sup>1</sup>, em parceria com a Academia do Instituto Nacional de Propriedade Industrial (INPI). A abordagem aqui adotada é a descrita em [Magerman et al. 2006]. Para a implementação, foi utilizado o Apache Lucene<sup>TM</sup>, uma biblioteca para indexação e busca *full-text* de documentos textuais, desenvolvida pela Apache Software Foundation.

## 2. Trabalhos Relacionados

Em seu trabalho em 2007, Tseng, Lin e Lin, argumentam sobre a importância estratégica da análise de patentes, enfatizando o esforço e o grau de especialização necessários para obter bons resultados. Neste sentido, reforçam a grande ajuda das técnicas automáticas nesse processo. Segundo os autores, os itens de análise mais críticos são aqueles que lidam com informações não estruturadas, em forma de textos. Eles propõem uma metodologia que utiliza técnicas de mineração de textos para análise *full-text* de patentes, consistindo de algumas etapas: pré-processamento dos documentos (criação da coleção, segmentação, parsing, etc.); indexação (extração de palavras-chaves, análise morfológica, filtragem de *stopwords*, associação de termos, etc.); agrupamento por tópicos (seleção de termos, categorização, etc.); e mapeamento (de tendências, agregação, etc.)([Tseng et al. 2007]).

Magerman, Looy e Song, em 2006, acrescentavam ainda os esforços no desenvolvimento de técnicas que produzem também a harmonização de nomes de entidades. Segundo eles, isso gera qualidade em termos de completeza, através da capacidade de capturar todas as variantes de nome de uma mesma patente e também em termos de acurácia, no que se refere à correta alocação de variantes para um único e já harmonizado nome. Eles desenvolveram um método de harmonização, no qual é acrescentado ao pré-processamento, a condensação de nomes, tendo conseguido reduzir em 17% o número de depositantes das bases testadas, nomes esses que correspondiam a entidades já lá representadas [Magerman et al. 2006]. A proposta aqui é testar essa metodologia, para que, caso viável, seja adicionada ao Periscope.

## 3. Harmonização dos nomes

O processo de harmonização faz com que nomes escritos de formas diferentes, que indicam uma mesma entidade, sejam considerados os mesmos. Isso gera um reagrupamento de documentos de patentes.

A versão atual do Periscope já faz a harmonização dos nomes, identificando e apontando algumas sugestões [Viebrantz 2013]. O usuário recebe essa lista de sugestões e então valida manualmente, selecionando aquelas aptas serem atualizadas no banco de dados. O intuito da utilização da abordagem aqui descrita é melhorar as sugestões apresentadas pelo software, trazendo um tratamento de nomes comuns e aplicando técnicas de busca por similaridade. Para a implementação, foi escolhido o Apache Lucene<sup>TM</sup>, que possui várias bibliotecas de tratamento de documentos textuais, que fazem desde o pré-processamento (retirada de caracteres acentuados, normalização, remoção das *stopwords*<sup>2</sup>, etc.), a indexação, até a busca, utilizando caracteres curingas, lógica booleana,

<sup>1</sup>Com o apoio da UFMT, do CNPq e da FAPEMAT

<sup>2</sup>Palavras que não contém valor útil para a aplicação em foco

medidas de similaridade, casamentos de strings [Navarro 2001], etc. Além das classes já disponíveis no Lucene, outras duas novas classes foram criadas, uma vez que sua sequência de processamento era crucial para a qualidade dos resultados: uma para a remoção dos descritores comuns e palavras indicativas de forma legal; a outra, necessária para a condensação do nome limpo.

As subseções a seguir descrevem as etapas de processamento aqui realizadas.

### 3.1. Pré-processamento

Para que haja eficiência no tratamento automático dos nomes, eles devem passar por limpeza, etapa da mineração de textos chamada de pré-processamento. A metodologia aqui utilizada prevê essas seguintes etapas:

- **Normalização dos caracteres do nome**

Para evitar que a caixa da letra atrapalhe as comparações, uma normalização nos caracteres do nome é necessária. O Lucene já disponibiliza este filtro, que passa todos os caracteres para caixa baixa (minúscula). Exemplo: `PEUGEOT AND CITROËN AUTOMOBILES SA` -> `peugeot and citroën automobiles sa`

- **Remoção de acentuação**

Caracteres acentuados atrapalham os resultados, pois possuem valores diferentes de seus equivalentes não acentuados. O Lucene possui uma implementação que passa todos os caracteres para o seu equivalente em ASCII. Exemplo: `peugeot and citroën automobiles sa` -> `peugeot and citroen automobiles sa`

- **Remoção de *stopwords***

*Stopwords* são palavras que ocorrem com muita frequência nos textos, podendo por este motivo atrapalhar o resultado da busca. Geralmente são utilizados para interligar palavras de uma frase, dando sentido e fluidez à leitura. Em alguns tipos de buscas, a retirada destas palavras pode alterar o valor ou o sentido do texto. Porém, com os nomes aqui trabalhados, não foi o caso. Este processo é realizado com os caracteres já normalizados, e sem acentuação, e já é implementado pelo Lucene. Exemplo: `peugeot and citroen automobiles sa` -> `peugeot citroen automobiles sa`

- **Tratamento das palavras de indicação de forma legal**

Palavras que indicam a forma legal de uma empresa (Inc, GMBH, Limited, etc.) acontecem com frequência em nomes de companhias, e em alguns casos, sua remoção não produz efeitos. Para evitar perda de informação, nos testes aqui realizados, apenas os indicativos de forma legal que aparecem no final dos nomes é que foram retirados. Exemplo: `peugeot citroen automobiles sa` -> `peugeot citroen automobiles`

- **Condensação do nome e geração do acrônimo**

Caracteres de separação como espaços, tabulações, traços, underlines, vírgulas, etc. devem também ser retirados. É essa a última etapa do processo. Esse processo de condensação faz com que o nome se torne bloco único, representado por apenas um *token*. Exemplo: `peugeot citroen automobiles` -> `peugeotcitroenautomobiles`.

No entanto, antes de remover os caracteres de separação, foi criado um acrônimo

com as iniciais de cada palavra do nome. Exemplo: peugeot citroen automobiles -> pca peugeotcitroenautomobiles.

### 3.2. Indexação

Neste processo, tanto o *token* condensado quanto seu acrônimo são indexados, mas individualmente, gerando um *token* com acrônimo e um *token* com o nome condensado.

### 3.3. Busca

Após o processo de indexação descrito acima, são realizadas então a busca por similaridade no nome condensado e ainda uma busca por caractere coringa no acrônimo. Os resultados obtidos são então mostrados ao usuário para ele faça a validação.

#### 3.3.1. Medida de similaridade com Damerau-Levenshtein

A técnica utilizada pelo Lucene para a busca de similaridade é a distância de Damerau-Levenshtein [Mccandless et al. 2010]. Para um padrão  $P$  de tamanho  $m$  e um texto  $T$  de tamanho  $n$ , a distância de Levenshtein destes textos é dada pela quantidade de inserções, remoções, substituições e transposições para transformar o padrão  $P$  em  $T$  ou vice-versa. Para o cálculo da distância, utilizou-se uma tabela  $D$ , de tamanho  $(m+1) \times (n+1)$ , onde, para cada posição da tabela  $D(i, j)$  temos o valor da distância  $E_L(P_{1..i}, T_{1..j})$  para  $i = 0, \dots, m$  e  $j = 0, \dots, n$ . A Tabela 1 mostra o cálculo de distância entre as palavras *bat* e *cat*. Verifica-se que a distância entre os dois é de 1, que é apenas a operação de substituição do caractere *b* e *c*.

Table 1. Tabela de distância entra as palavras *bat* e *cat*

|   |   |   |   |
|---|---|---|---|
|   | c | a | t |
| 0 | 1 | 2 | 3 |
| b | 1 | 1 | 2 |
| a | 2 | 1 | 1 |
| t | 3 | 1 | 1 |

Em suas versões mais recentes (a partir da 4.0), o Lucene passou a adotar a estrutura de um Autômato Finito Determinístico para realizar suas buscas, o que trouxe uma melhora significativa em relação ao tempo consumido. O Autômato de Levenshtein utilizado por ele é descrito por Schulz em [Schulz and Mihov 2002]; dado um texto de entrada  $P$ , é gerado um autômato  $A_P$  que aceite um número máximo  $n$  de distância com os textos já indexados.

## 4. Resultados

Com uma base de 40263 nomes de depositantes diferentes, foram feitos testes de indexação de dois modos diferentes: com a nova abordagem, aqui descrita, e com a antiga abordagem, já existente no Periscope. Dos nomes iniciais, foram retirados aleatoriamente 6 nomes, que foram utilizados na busca sobre os índices gerados. Os resultados são mostrados na Tabela 2.

**Table 2. Resultado da busca com a nova e antiga metodologia**

| Nome  | Nova metodologia   | Antiga metodologia   |
|---|--|--|
| BAYER AG                                    | BAYER AG<br>BAYER IP GMBH<br>BAYER SAS<br>BAYER PHARMACEUTICALS CORP<br>BAYER NEW ZEALAND LTD  | BAYER AG<br>BAYER SCHERING AG<br>BAYER MATERIALSCIENCE AG<br>BAYER HEALTHCARE AG<br>BAYER PHARMA AG<br>BAYER MATERIALSCIENCE AG<br>BAYER CROPSCIENCE AG<br>BAYER SCHERING PHARMA AG  |
| BAYER ANIMAL HEALTH GMBH                    | BAYER ANIMLA HEALTH GMBH<br>BAYER ANIMAL HEALTH GMBH<br>BAHL PARAMVIR<br>BAHADUR PRASHANT<br>BAHUKUDUMBI PRADIPKUMAR<br>BAHADUR KHARKATHOKI PURNA<br>BAHN PATRICK R E<br>BAHAROV GEORGI<br>BORUTA ANDRZEJ B HANDLOWE KINESIS | BAYER ANIMAL HEALTH GMBH<br>BAYER ANIMLA HEALTH GMBH<br>ELANCO ANIMAL HEALTH IRELAND LTD<br>BAYER IP GMBH<br>BAYER INNOVATION GMBH<br>BAYER TECHNOLOGY SERVICES GMBH<br>Bayern-Chemie Gesellschaft für flugchemische Antriebe mbH<br>DBK DAVID   |
| BALLY GAMING INC                            | BALLY GAMING INC<br>BALLY GAMING INT INC<br>BALLY GAMING   | BAADER GMBH<br>BALLY GAMING INC<br>BALLY GAMING INT INC<br>BALLY GAMING<br>LAMINA INC<br>KONAMI GAMING INC<br>WMG GAMING INC   |
| KING JIM CO LTD                             | KING JIM CO LTD<br>KING COM LTD<br>KING I TECH CORP  | KING JIM CO LTD<br>KING COM LTD<br>JVM CO LTD<br>KING SLIDE WORKS CO LTD<br>KING BILLION ELECTRONICS CO LTD<br>LTD KING SLIDE WORKS CO<br>SUNNY KING MACHINERY CO LTD<br>CHUNG KING ENTPR CO LTD<br>KING YUAN ELECTRONICS CO LTD   |
| PHILIPS DEUTSCHLAND GMBH                    | PHILIPS DEUTSCHLAND GMBH<br>PHILIPS ELECTRONICS NV<br>PHILIPS ELECTRONICS NA<br>R A PHILIPS IND INC<br>KONINK PHILIP N V<br>PHILIPS NV<br>KONINKLIJKE PHILIPS N.V<br>PHILIPS INTERNAT B V INTELLECTUAL PROPERTY & STANDARDS  | PHILIPS DEUTSCHLAND GMBH<br>SANOFI AVENTIS DEUTSCHLAND GMBH<br>AIRBUS DEUTSCHLAND GMBH<br>PHILIPS ELECTRO SPECIAL GMBH<br>ISTA DEUTSCHLAND GMBH<br>EADS Deutschland GmbH<br>Eurocopter Deutschland GmbH<br>Airbus Deutschland GmbH<br>EADS DEUTSCHLAND GMBH<br>INVACARE DEUTSCHLAND GMBH   |
| FUJI ELECTRIC CO.                           | FUJI ELECTRONIC IND CO LTD<br>FUJI ELECTRIC CO LTD<br>FUJI CHEMICALS INDUSTRIAL CO.<br>FUJI CHEM IND CO LTD<br>FUJI ELECTRIC CO.<br>FUJI CHEMICAL INDUSTRY CO.<br>FORSCHUNGSGESELLSCHAFT FUER AR<br>Furui<br>WU YI           | FUJI ELECTRIC CO LTD<br>FUJI ELECTRIC CO.<br>FUNAI ELECTRIC CO<br>FUJI ELECTRONIC IND CO LTD<br>FUJI ELECTRIC RETAIL SYSTEMS<br>FUJI ELECTRIC HOLDINGS<br>FUJIAN ERLING ELECTRONIC CO LTD<br>FUJIAN FUDA BIOTECH CO LTD  |
| AMERICAN STERILIZER CO                      | AMERICAN STERILIZER COMPANY  | AMERICAN STERILIZER COMPANY<br>AMERICAN NATIONAL CAN CO<br>BRITISH AMERICAN TOBACCO CO<br>AMERICAN CYANAMID CO<br>HOECHST CO AMERICAN<br>AMERICAN DISTRICT TELEGRAPH CO<br>BRITISH AMERICAN TOBACCO CO<br>AMERICAN MOLDED PRODUCTS CO<br>AMERICAN POWER CONY CORP  |
| ARCELORMITTAL INVESTIGACION Y DESARROLLO SL | ARCELORMITTAL INVESTIGACION Y DESARROLLO S L<br>AIDEE PATRICK<br>AID SA<br>AIDC AEROSPACE IND DEV CORP<br>ARCELORMITTAL INVEST Y DESARROLLO<br>ACUMENER INVESTIGACION Y DESARROLLO SL  | ARCELORMITTAL INVEST Y DESARROLLO S L<br>ARCELORMITTAL INVEST Y DESARROLLO<br>ARCELORMITTAL INVESTIGACION Y<br>FUNDACION AGENCIA ARAGONESA PARA LA INVESTIGACION Y EL DESARROLLO<br>CT DE INVESTIGACION Y DESARROLLO DE MEDICAMENTOS<br>INICIA INGENIERIA INVESTIGACION Y DESARROLLO S L<br>GERDAU INVESTIGACION Y DESARROLLO EUROPA S A |
| GAZPROMNEFT NOJABR SKNEFTEGAZ               | GAZPROMNEFT NOJABR SKNEFTEGAZG<br>GIVAUDAN NEDERLAND SERVICES B V  | GAZPROMNEFT NOJABR SKNEFTEGAZG<br>OOO NEFTEGAZ STAL EHKSPERTNO N<br>NOAR J BARRY   |
| CONTINENTAL TEVES AG & CO OHG               | Continental Teves AG & Co. OHG   | Continental Teves AG & Co. OHG<br>KREUZER GMBH + CO. OHG<br>SEVEX AG<br>OG CORP<br>AIR BAG PACKING CO LTD<br>BAG CORP<br>OHI SEISAKUSHO CO.<br>OC OERLIKON BALZERS AG  |

Uma análise superficial preliminar apontou resultados ainda aquém dos esperados em alguns cenários. Por exemplo, ao examinar os resultados retornados pela busca à consulta “BAYER AG”, é possível notar variações relacionadas ao termo “BAYER”; porém a nova metodologia falhou em retornar todas as variações. A falha nesses resultados pode ser explicada pelos descritores comuns utilizados. A metodologia aqui seguida foi a de utilizar uma lista encontrada em [Magerman et al. 2006], no qual as palavras MATERIALSCIENCE, CROPSCIENCE e SCHERING não se encontravam na lista. Essa ausência levou a um resultado impreciso. Esta questão será discutida na Seção 5.

Já ao analisar a consulta “BALLY GAMING INC”, é possível perceber que a nova metodologia se aplica de uma melhor forma. Esse resultado se dá pela retirada os descritores comuns e palavras que causam variações aos nomes. A retirada desses elementos corrigiu os resultados falhos da antiga implementação. Isso é causado porque os termos “GAMING” e “INC” possuem pesos altos no nome, e portanto, influenciam no retorno da busca.

## 5. Conclusões

A metodologia descrita neste artigo mostrou ser promissora. Juntamente com a implementação utilizando o Apache Lucene™, uma biblioteca bastante completa para o tratamento de documentos textuais. O problema da harmonização de nomes ainda é um dos gargalos na área de análise de patentes, justamente por se deparar com problemas difíceis de serem tratados.

Tendo agora o domínio da tecnologia, há várias questões a serem trabalhadas. Torna-se também necessária a presença de um especialista do domínio para definir as harmonizações possíveis e para validar os resultados produzidos pela ferramenta.

A análise dos descritores comuns provou ser uma das etapas mais importantes. A produção de uma lista mais representativa desses descritores poderá trazer melhoria nos resultados finais. Este será o próximo passo a ser realizado neste trabalho. Deverá haver uma maior lapidação das técnicas de aproximação aqui utilizadas e também do estudo de novas técnicas que poderão ser confrontadas e até combinadas. A busca com acrônimos precisa também ser melhor explorada, com resultados combinados à busca por similaridade. A utilização de etiquetas morfossintáticas também pode ser considerada.

## References

- Abbas, A., Zhang, L., and Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information*, 37(0):3 – 13.
- do Prado, H. A. and Ferneda, E. (2008). *Emerging Technologies of Text Mining: Techniques and Applications*. Information Science Reference - IGI Global, New York, 1 edition.
- Magerman, T., Looy, B. V., and Song, X. (2006). Data production methods for harmonised patent statistics : Patentee name harmonisation. *EUROSTAT Working Paper and Studies*, 1(1):85.
- Mccandless, M., Hatcher, E., and Gospodnetic, O. (2010). *Lucene In action*. Manning Publications Co., second edition.
- Navarro, G. (2001). A Guided Tour to Approximate String Matching. *ACM Computing Surveys*, 33(1):31–88.
- Schulz, K. U. and Mihov, S. (2002). Fast string correction with Levenshtein automata. *IJDAR - International Journal on Document Analysis and Recognition*, 5(1):67–85.
- Tseng, Y.-H., Lin, C.-J., and Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information Processing and Management*, 43(5):1216 – 1247. Patent Processing.
- Viebrantz, A. F. P. M. (2013). *Bases de Patentes: Abordagem utilizando NoSQL e Full-Text*. Trabalho de conclusão de curso, Universidade Federal de Mato Grosso.

## Indicadores em um PMO de Projetos de Pesquisa

**Maykon P. dos Reis, Diego F. Brock, Paulo Henrique Garcia, Josefson F. de Souza,  
Nilton H. Takagi, Josiel M. Figueiredo**

Instituto de Computação - Universidade Federal de Mato Grosso (UFMT)

Av. Fernando Corrêa da Costa, nº 2367 - Bairro Boa Esperança. Cuiabá - MT - 78060-900

maykon.net1@gmail.com, diegobrock@gmail.com, phgarcia.mt@gmail.com  
josefson.br@gmail.com, nilton@ic.ufmt.br, josiel@ic.ufmt.br

**Abstract.** *For managing, indicators are necessary to measure the progress of activities and it is not different for projects. Every project has one or more deliverable results, considering it as a temporary endeavor, it involves material and human resources and requires previous budget and financial management of investments. For managing all these variables, indicators are essential. Considering indicators of management of research projects and other academic projects, four indicators to measure the progress of scope, time and cost of the projects were implemented. The indicators were created based on the PMBOK and on the expertise of those involved, considering the most essential GAPS for the control and monitoring of the project portfolio.*

**Keywords.** *Project management; indicators; Research Projects; Monitoring and control.*

**Resumo.** *Para gerenciar, os indicadores são necessários para aferir o andamento das atividades e com projetos não é diferente. Todo projeto tem um ou vários parâmetros de avaliação do resultado final, sendo um empreendimento temporário, envolve recursos materiais, humanos e exige orçamento prévio e gestão financeira dos investimentos realizados. Para gerenciar todas essas variáveis, os indicadores são essenciais. Considerando indicadores de gestão de projetos de pesquisa e demais projetos acadêmicos, foram implementados quatro indicadores para medir o andamento do escopo, tempo e custo dos projetos. Os indicadores criados foram baseados no PMBOK e na expertise dos envolvidos, considerando os GAPS mais essenciais para o controle e monitoramento do portfólio de projetos.*

**Palavras-chave.** *Gerenciamento de Projetos; Indicadores; Projetos de Pesquisa; Monitoramento e controle.*

## 1. INTRODUÇÃO

Indicadores de projetos são válidos para todos os gerentes que precisam planejar algum tipo de programa ou projeto, para a execução de ações, para a análise de projetos ou para o gerenciamento dos processos de trabalho, onde será possível tomar decisões para mudar o rumo do projeto ou mesmo para paralisá-lo, sendo função dos indicadores mostrar da forma mais clara possível a situação atual e as tendências do projeto.

A devida utilização de indicadores de desempenho no gerenciamento de projetos se torna indispensável para o efetivo acompanhamento e tomada de decisões. A não-utilização seria o mesmo que monitorar a febre de uma pessoa sem utilizar um termômetro, usando apenas o contato físico ou a aparência da pessoa para efetuar a avaliação. Naturalmente que pode funcionar mesmo com baixo nível de precisão, mas também pode falhar, com impacto gravíssimo (BARCAUI, 2012).

Segundo Ferreira, Cassiolato e Gonzalez (2009), a utilização de indicadores é uma forma de medida de ordem quantitativa ou qualitativa, de significado particular e utilizada para organizar e captar as informações relevantes dos elementos que compõem o objeto da observação. Pode ser definida também, como um recurso metodológico que informa empiricamente sobre a evolução do aspecto observado.

O grande desafio é traduzir essa definição em indicadores relevantes, pensar que se algo é importante, deverá ser detectável. E se for possível detectar, logo poderá ser medido. Onde, se for possível medir, poderá ser monitorado e controlado.

O objetivo dos indicadores é fornecer informação atualizada, oportuna e em tempo real para facilitar a tomada de decisões. O desafio maior da abordagem deste projeto é relacionar esses indicadores em um ambiente de pesquisa, no qual a incerteza é algo intrínseco.

Este artigo trata da aplicação de indicadores em um ambiente universitário no qual projetos diversos são realizados. O enfoque principal aqui é de tratar de projetos de pesquisa com o uso de indicadores genéricos de desempenho, visando atender aos resultados esperados que representam as propriedades comuns de projetos diversos.

## 2. O AMBIENTE DE PESQUISA: PMO-IC

O *Project Management Office* (PMO) ou Escritório de Gerenciamento de Projetos é definido como a unidade organizacional com o objetivo de conduzir, planejar, organizar, controlar e finalizar as atividades dos projetos, possuindo pessoas com conhecimentos em Gerenciamento de Projetos, capazes de prestarem todo o suporte necessário aos gerentes de projeto e sua equipe. Muitos o denominam de QG (Quartel General), pois é o centro de informações de controle (PMI, 2013).

Um PMO torna-se necessário para integrar e apoiar as atividades de Gerenciamento de Projetos, analisar dados, auxiliar na tomada de decisão e, por consequência, retirar a sobrecarga dos gerentes de projetos. Contribui com inúmeras melhorias no Gerenciamento de Projetos. Torna-se um centro de apoio onde as informações encontram-se centralizadas e auxiliam no suporte a decisões estratégicas. Monitora todos os prazos, orçamentos, progresso dos projetos e coordena os padrões de qualidade, disponibilizando e mantendo uma metodologia de Gerenciamento de Projetos na organização. Neste cenário o PMO-IC está sendo estruturado de modo a possuir maior auxílio no acompanhamento dos projetos de pesquisa relacionados ao Instituto de Computação (IC-UFMT), e irá utilizar os registros dos projetos e suas respectivas atividades cadastradas em um Sistema de Gerenciamento de Projetos (SIGP), sendo atualmente utilizado o GP-WEB.

O GP-WEB é um sistema integrado de gerenciamento de projetos que utiliza como base as práticas de gestão de projetos para catálogo dos registros de projetos e atividades nele

inserido. As funcionalidades de gerenciamento de projetos estão alinhadas a Metodologia de Gerenciamento de Projetos do Ministério do Planejamento (MGP-SISP) e contempla os processos definidos na metodologia do PMO-IC. Está disponível via plataforma Web, desenvolvido em linguagem PHP e utiliza com SGDB o MySQL.

### **3. DEFINIÇÃO DOS INDICADORES**

Os indicadores do projeto medem os benefícios esperados, onde normalmente aparecem depois que o projeto foi finalizado. Já os indicadores da gestão do projeto estão relacionados ao controle durante a execução do projeto, onde as entregas previstas são comparadas com as entregas realizadas.

De acordo com Mendonça (2002), indicadores são normalmente constituídos por duas unidades de medida correlacionadas. Servem para medir o desempenho de uma determinada atividade ou rotina, observando se o que está sendo feito está dentro dos parâmetros exigidos, ou seja, se está de acordo com o que foi planejado.

De acordo com Vieira (1999), quem irá executar o processo deverá fazer a medição, pois quem está nesta função normalmente terá conhecimento dos pontos críticos e os GAPs da rotina, devendo produzir os indicadores para apurar aquilo que está acontecendo no processo. Não se deve colocar uma pessoa para medir o que a outra faz, pois assim onera-se a organização.

Deverá ser definido pelo Gestor do Projeto as unidades de medida fundamentais para o gerenciamento do projeto e trabalhar sua conceituação, sempre mantendo o cuidado para garantir o entendimento por todos os interessados da ação. O indicador deve ser montado com nome, objetivo, fórmula, periodicidade e fonte. É importante não dar o mesmo nome para indicadores diferentes na organização (MENDONÇA, 2002).

Após definidos os indicadores é necessário realizar a devida implementação estabelecendo plano de metas, definindo percentuais ou quantitativos a serem atingidos em determinado espaço de tempo. Deve-se acompanhar os resultados e representar graficamente os índices, demonstrando a evolução do indicador ao longo do tempo. Isso deve ser definido no planejamento do projeto.

Para o ambiente em questão, foi utilizado os indicadores conforme a Tabela 1.

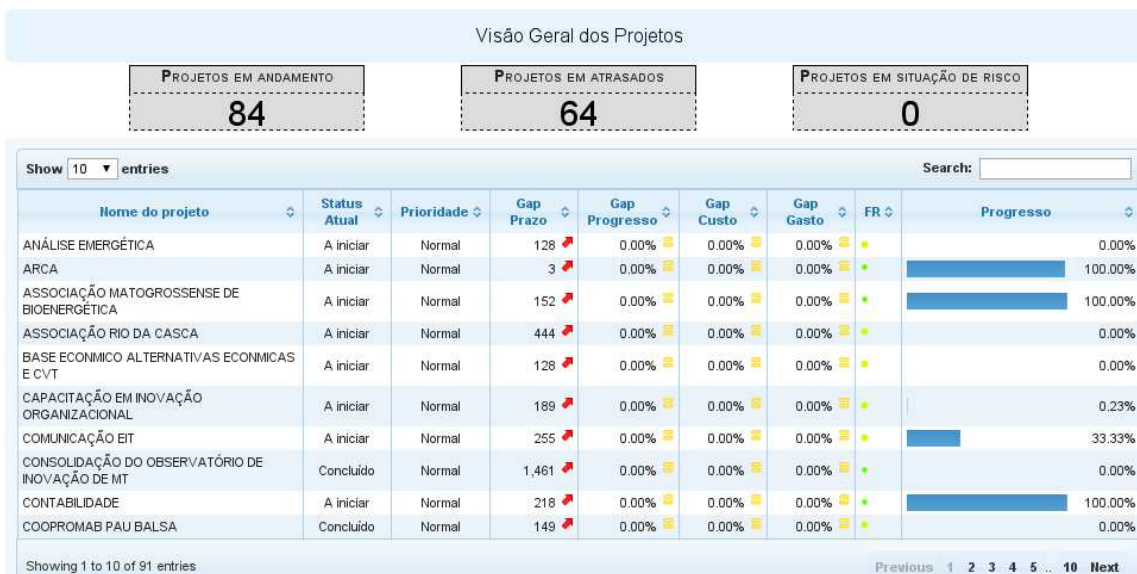
Tabela 1 - Descrição dos Gaps elaborados

| Gap                     | Objetivo  | Variáveis   | Avaliação  |
|-------------------------|---|---|--|
| <b>Gap de Prazo</b>     | Medir o distanciamento entre o prazo registrado na baseline mais recente do projeto e seu prazo real.         | GP: Gap de Prazo;<br>PB: Prazo em dias, registrado na baseline;<br>PR: Prazo em dias real;<br>Temos que $GP = PR - PB$ ;  | Considera-se como um bom resultado quando $GP \leq 0$ .<br>Se $GP > 0$ então considera-se o resultado como Ruim;   |
| <b>Gap de progresso</b> | Medir o distanciamento entre o progresso registrado na baseline mais recente do projeto e seu progresso real. | GPR: Gap de Progresso;<br>PRB: Progresso registrado na última baseline do projeto;<br>PRR: Progresso real do projeto;<br>Temos que: $GPR = ((PRR * 100) / PRB) - 100$ | Considera-se como um bom resultado quando $GPR \geq 0$ .<br>Se $GPR < 0$ então considera-se o resultado como Ruim;   |
| <b>Gap de custo</b>     | Medir o distanciamento entre o custo registrado na baseline mais recente do projeto e seu custo real.         | GC: Gap de Custo;<br>CB: Custo registrado na última baseline do projeto<br>CP: Custo real do projeto<br>Temos que: $GC = ((CP * 100) / CB) - 100$                     | Considera-se como um bom resultado quando $GC \leq 0$ .<br>Se $GC > 0$ então considera-se o resultado como Ruim; O custo de um projeto no Gp-web sempre é o que está planejado para se gastar de fato.   |
| <b>Gap de Gasto</b>     | Medir o distanciamento entre o gasto registrado na baseline mais recente do projeto e seu gasto real.         | GG: Gap de gasto;<br>GB: Gasto registrado na última baseline;<br>GR: Gasto real do projeto;<br>Temos que: $GG = ((GR * 100) / GB) - 100$                              | Considera-se como um bom resultado quando $GG \leq 0$ .<br>Se $GG > 0$ então considera-se o resultado como Ruim; O gasto no gp-web é sempre o que de fato foi gasto, podendo ser o mesmo valor do custo planejado, ou também acima ou abaixo do mesmo. |

#### 4. OS RESULTADOS

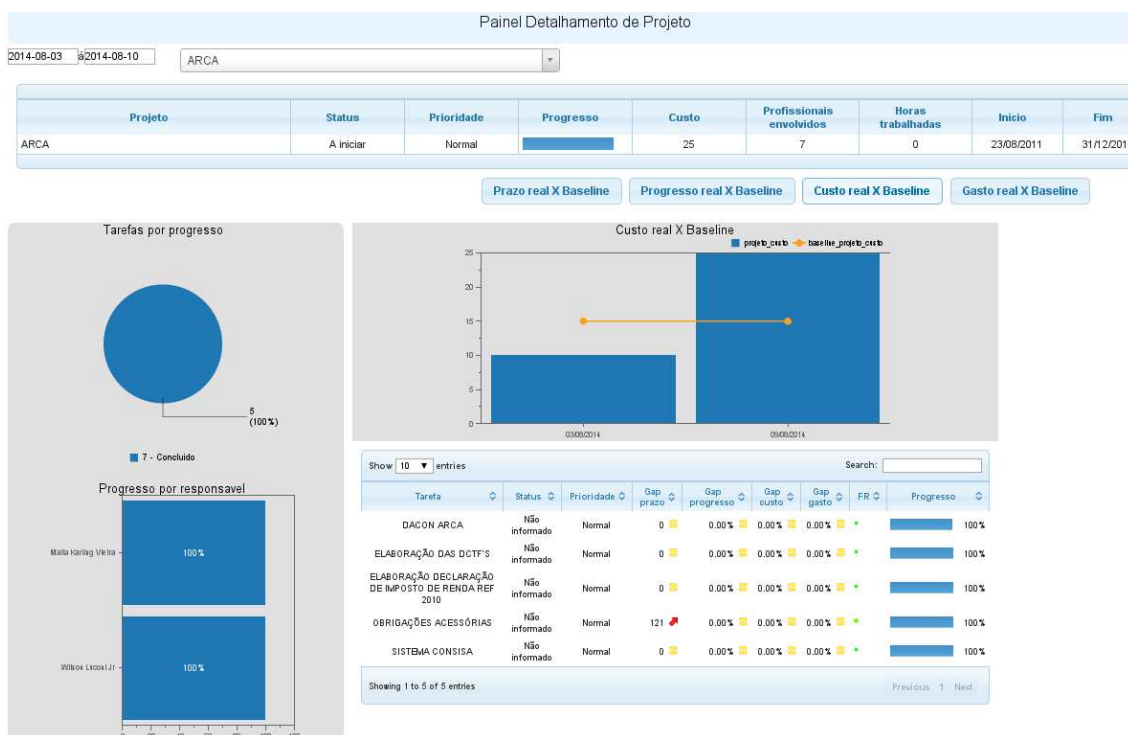
Foram desenvolvidas análises de baselines após o processo de tratamento da base de dados do SIGP, neste caso o sistema GP-WEB. Análises nas estruturas transacionais do SIGP foram realizadas, seguindo para a extração, tratamento e carga dos dados com base no modelo multidimensional desenvolvido e visões de indicadores conforme os requisitos levantados. Sendo assim, todo o processo de *Business Intelligence* (BI) foi realizado para demonstração dos dados em forma gráfica nos dashboards conforme demonstrado abaixo.

Foram desenvolvidas na ferramenta *Pentaho Analytics* os seguintes *dashboards* conforme Figura 1.



**Figura 1 - Dashboard como a Visão geral dos Projetos da Organização**

Nesta visão são listados os projetos da organização sendo informados o *status* atual de cada projeto, a prioridade, e as variáveis de prazo, progresso, custo e gasto de acordo com o comparativo da situação real do projeto com as baselines. A partir destas variáveis é gerado a contabilização do fator de Risco (FR) de cada projeto. Ao clicar em um projeto da listagem, o usuário será redirecionado para o *dashboard* com o detalhamento do projeto conforme a Figura 2.



**Figura 2 - Dashboard com o detalhamento do projeto selecionado**

Já nesta visão detalhada do projeto é possível visualizar os indicadores das tarefas do projeto selecionado, sendo possível selecionar quais das baselines deseja realizar a comparação. É exibido também o detalhamento de todas as tarefas do projeto, sendo realizada a mesma análise das variáveis de status atual de cada tarefa, a prioridade, e as variáveis de

prazo, progresso, custo e gasto de acordo com as baselines. A partir destas variáveis é gerado a contabilização do fator de Risco(FR) de cada tarefa. O Fator de risco é resultante numa escala de 1 a 5, a variação não esperada dos 4 GAPs, além da situação da tarefa (atrasada ou não). O status é sinalizado como "Em risco" quando o resultante da soma dos fatores de risco for maior ou igual a 3.

## 5. CONCLUSÃO

Possuir a capacidade de análise de cada informação de um projeto é sem dúvida um fator primordial para se chegar a excelência do resultado esperado, ou o mais próximo do planejado. Utilizar recursos para realizar o gerenciamento de projetos se torna necessário quando possuímos uma quantidade de informações que dificilmente seriam organizadas de modo a extrair informações de forma manual. É de suma importância a utilização de uma ferramenta que possa prover informações com os dados necessários para diversas análises. A utilização de processos automatizados de indicadores gera economia de tempo e custo operacional, onde tais informações podem ser levantadas em pouco tempo dentro da ferramenta de BI e exibidas de forma dinâmica no módulo de análise de dashboard.

É importante ressaltar que procurou-se utilizar indicadores cuja obtenção seja feita de forma automática, sem a necessidade de esforço adicional para obtenção dos mesmos.

Também se faz necessário a utilização e divulgação da metodologia de gerenciamento de projeto na organização, para que as informações geradas no sistema de gerenciamento sejam inseridas de maneira correta, possibilitando assim uma base de informações catalogadas consistente, sem deixar que determinadas análises se tornem inviáveis pela falta de determinadas informações.

Como trabalho futuro, está planejado a inclusão de indicadores relacionados a metas dos projetos, como quantidade publicações realizadas e quantidade de pessoas atingidas pelos projetos, sendo estes, indicadores de resultado fundamentais para agregação dos pesquisadores ao processo.

## 6. REFERÊNCIAS BIBLIOGRÁFICAS:

BARCAUI, André B.; **PMO: escritório de projetos, programas e portfólio na prática**. Rio de Janeiro: Brasport, 2012.

FERREIRA, H.; CASSIOLATO, M.; GONZALEZ, R. **Uma experiência de desenvolvimento metodológico para avaliação de programas: o modelo lógico do programa segundo tempo. Texto para discussão**. Rio de Janeiro: IPEA, 2009.

MENDONÇA, Mauro. **TAM – Técnicas para Análise e Melhoria de Processos**. Curso em Fita de Vídeo, Link Quality, 2002.

PENTAHO. **Pentaho Business Analytics. Pentaho**. 2005-2014. Disponível em <<http://www.pentaho.com/product/business-visualization-analytics>>. Acesso em: jul. 2014.

PMI – PROJECT MANAGEMENT INSTITUTE. **A guide to the project management body of knowledge**. 5.ed. Pennsylvania: PMI, 2013.

VIEIRA, Sonia. **Estatística para a Qualidade – Como Avaliar com Precisão a Qualidade em Produtos e Serviços**. Rio de Janeiro: Ed. Campus, 1999.

# Análise de simulação de protocolo de roteamento geográfico para VANET

**Pedro Clarindo da Silva Neto, Rafaely Schwaab Sampaio, Cheiene Batista Oliveira**

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT)– Campus Pontes e Lacerda

Rodovia MT-473, s/n - CEP: 78250-000– Pontes e Lacerda– MT – Brasil

pedro.neto@plc.ifmt.edu.br, rafa\_22p1@hotmail.com, cheiene@msn.com

***Abstract.** This paper approaches the vehicular ad hoc network (VANET) routing, more precisely, geographic routing, and parse through a module of the NS-3 simulator, the Greedy Perimeter Stateless Routing protocol (GPSR), to verify its effectiveness and the protocol behavior in a scenario of moving nodes.*

***Resumo.** O presente trabalho tem como foco abordar o roteamento em redes veiculares ad hoc - VANET, mais precisamente, roteamento geográfico, e, analisar através de um módulo do simulador NS-3, o protocolo Greedy Perimeter Stateless Routing (GPSR) para verificar sua efetividade e o comportamento do protocolo em um cenário de movimentação de nós.*

## 1. Introdução

Em virtude das características da mobilidade de seus nós, uma rede pode ser classificada como infraestruturada ou em modo ad hoc. No primeiro caso, toda a comunicação entre dispositivos móveis da rede ocorre através de estações bases fixas, por exemplo, pontos de acesso em redes IEEE 802.11. No caso da rede em modo ad hoc, a rede é composta por dispositivos móveis com capacidade para comunicarem entre si, não havendo necessidade de se recorrer a uma infraestrutura pré-existente pois são utilizados os nós intermediários. Nessas redes existem problemas no encaminhamento de pacotes, uma vez que os diferentes nós não têm uma localização definida devido à sua mobilidade.

Redes ad-hoc móveis (MANETs) fornecem comunicação peer-to-peer e multi-hop em uma rede de auto-configuração de dispositivos móveis. MANETs utilizam protocolos diferentes, dependendo da mobilidade da rede. As redes *ad hoc* veiculares (VANETs) são um dos tipos de MANET, em que os nós móveis (neste caso, veículos) alcançam altas velocidades. Os nós da rede VANET estão sujeitos a períodos sem nenhum tipo de ligação entre veículos, em virtude do tráfego veicular ou a atrasos na recepção de pacotes, devido, geralmente, às velocidades elevadas dos veículos e do cenário em que estão inseridos (Coelho, 2011).

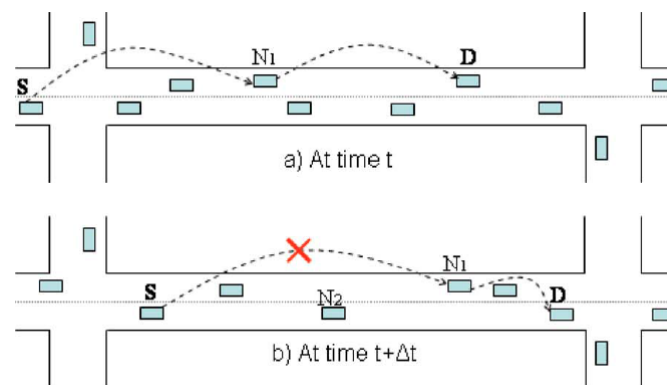
De acordo com Yousefi, Mousavi, & Fathy, (2006), VANET, é uma subclasse de redes MANETs. É uma abordagem promissora para o futuro dos sistemas de transportes inteligentes, que não possuem infraestrutura fixa e em vez disso contam com os próprios veículos para fornecer funcionalidade de rede. Para Li & Wang, (2007), VANET é uma nova tecnologia emergente que integra rede *ad hoc*, sem fio LAN (WLAN) e tecnologia de celular para obter comunicações inteligentes entre veículos e melhorar a segurança e eficiência da circulação rodoviária. VANETs distinguem-se de outros tipos de redes *ad hoc* pelas suas arquiteturas de rede híbrida, características do movimento de nós, e novos cenários de aplicação.

Contudo, as análises dos protocolos de roteamento tradicionais para redes móveis ad hoc (MANETs) demonstraram que o seu desempenho é pobre em cenários com VANETs. O principal problema com estes protocolos, por exemplo, AODV e DSR, em ambientes VANETs é a sua instabilidade de rota (Nzouonta, Rajgure, & Borcea, 2009). Uma outra abordagem são os protocolos de roteamento geográfico. Estes protocolos não estabelecem rotas, mas usam a posição do destino e a posição dos nós vizinhos para transmitir dados. Este trabalho visa analisar a implementação de um protocolo de roteamento geográfico para VANET desenvolvido para o simulador NS3 e verificar seu comportamento em um cenário simulado com e sem mobilidade de nós.

## 2. Roteamento em VANET

Avaliações experimentais têm demonstrado que protocolos de roteamento que não usam a localização geográfica, tais como protocolos baseados em topologia, não são escaláveis. Trabalhos de avaliação de desempenho mostraram que nem OLSR, nem AODV são capazes de fornecer taxa de entrega de pacotes aceitável em um cenário VANET (Jain, Puri, & Sengupta, 2001; Khan & Qayyum, 2009).

A tradicional visão das rotas é uma sucessão de nós fixos entre a origem e o destino que conduzem a frequentes rotas quebradas na presença da alta mobilidade das VANETs. Consequentemente, muitos pacotes são descartados, e há sobrecarga devido a reparos de rotas ou aumento significativo de notificação de falhas, levando a baixos índices de entrega e atrasos de transmissão elevadas (Fonseca, Camões, & Vazão, 2012; Nzouonta et al., 2009). Na Figura 1, observa-se o exemplo de uma rota quebrada devido a movimentação de um nó, nesse caso, um veículo.



**Figura 1. Rota (S, N1, D), que foi estabelecida no tempo t quebra no tempo t + Δt quando N1 se move para fora da faixa de transmissão de S (Nzouonta et al., 2009).**

Ao contrário do roteamento *node-centric*, o roteamento geográfico tem a vantagem de que qualquer nó que garanta o progresso para o destino pode ser usado para o encaminhamento (Nzouonta et al., 2009).

Segundo Fonseca, Camões, & Vazão (2012), os primeiros protocolos de roteamento baseados em posições baseavam-se no conceito de progresso para o nó de destino. Soluções melhoradas tem sido propostas, em que a seleção é baseada na distância geográfica para o nó de destino.

Nesta classe de protocolos, é sempre assumido que cada nó tenha informações sobre a posição disponível, que é uma suposição válida atualmente, uma vez que a maioria dos

carros tem um dispositivo de GPS. Para esses protocolos, o nó precisa saber a localização do destino, que é feito através de um serviço de localização, como a *Grid Location Service* (GLS), o *Reactive Location Service* (RLS) ou *Hierarchical Location Service* (HLS). Nesse caso, a localização do destino é feita no pacote para que os nós retransmitidos não precisem usar o serviço de localização de novo, reduzindo a sobrecarga.

### 3. GPSR – Greedy Perimeter Stateless Routing

Segundo Karp & Kung (2000), Greedy Perimeter Stateless Routing (GPSR) é um protocolo de roteamento para redes de datagramas sem fio, que utiliza as posições de roteadores e destino de um pacote para tomar decisões de encaminhamento de pacotes. GPSR toma decisões de encaminhamento gananciosos usando apenas informações sobre vizinhos imediatos de um roteador na topologia da rede. Quando um pacote alcança uma região, onde o encaminhamento ganancioso não é possível, o algoritmo recupera pelo roteamento em torno do perímetro da região.

Ao manter o estado apenas sobre a topologia local, GPSR escala melhor em estado *per-router* do que a abordagem de menor caminho e de roteamento *ad hoc* de protocolos, conforme o número de destinos da rede aumenta. Sob mudanças na topologia frequentes de mobilidade, GPSR pode usar informações de topologia local para descobrir rapidamente novas rotas corretas.

Nessa abordagem, um nó que quer enviar um pacote escolhe seu vizinho geograficamente mais perto do destino, como seu próximo salto. Um nó pode calcular a distância localmente, uma vez que mensagens Hello têm a posição dos nós e a posição do destino é fornecida por um serviço de localização. Em estratégias de encaminhamento ganancioso podem ocorrer situações chamadas de máximo local ou ótimo local, em que o veículo remetente está mais perto do destino do que todos os seus vizinhos, e o destino não está acessível por um salto.

Quando ocorre um máximo local, uma estratégia de recuperação é usada. A estratégia utilizada pelo GPSR é a regra da mão direita, comumente usada para atravessar grafos (Fonseca, Camões, & Vazão, 2012).

O algoritmo 1 descreve o GPSR usando um pseudo-código, onde:

- R, é o nó que recebe um pacote de p para Destino D;
- N, é o conjunto de vizinhos de um salto de R;
- n, é um nó do conjunto N, que é utilizado para transmitir o pacote;
- D, é o destino do pacote.

```

se  $\exists n \in N : Distancia(n,D) \leq Distancia(R,D)$  entao
{Encaminhamento Ganancioso}
 $n = MinDistancia(N,D)$ 
EncaminhaPacote(p,n)
Retorna
senao
{Local maximo, use regra mao direita}
 $n = RegraMaoDireita(p,n)$ 
EncaminhaPacote(p,n)
Retorna
fimse

```

**Algoritmo 1: Pseudo-código GPSR (Traduzido de Fonseca, Camões, & Vazão., 2012).**

Os principais componentes da implementação são: os serviços de localização, manutenção de vizinho e a seleção do próximo salto. Para atender às necessidades de um serviço de localização, foi implementado um serviço onisciente (God Location Service).

A implementação utilizada é baseada no código disponibilizado por (Fonseca, Camões, & Vazão, 2012), para o simulador Network Simulator versão 3 (NS3). Esta implementação foi escolhida devido ao fato de não haver implementações de protocolos de roteamento geográfico para o NS, que é o simulador de redes mais amplamente utilizado dentro dos trabalhos acadêmicos, além de possuir a peculiaridade de ter sido projetado para VANET. Foi realizada a implementação da geração do arquivo de rastreamento que possibilitou a visualização do protocolo em ação nos cenários simulados através do Network Animator, onde pode ser observado os saltos entre os nós, conforme Figura 2.

#### 4. Metodologia

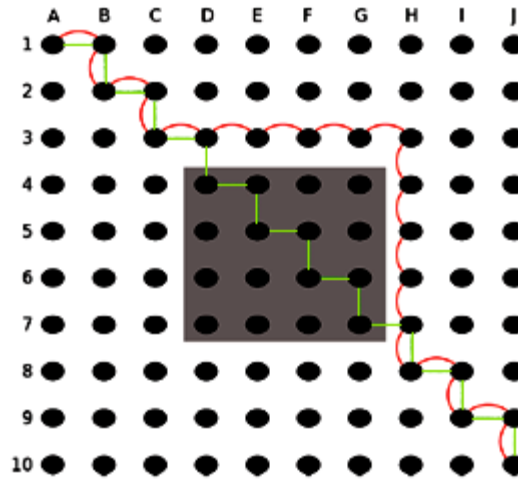
A topologia utilizada é composta por uma grade de nós de 10x10, separados em 100 metros. A faixa de transmissão foi ajustada para 100 m, de forma que nós na diagonal não são vizinhos. O MAC 802.11 DCF foi usado sem RTS / CTS. O modelo de propagação utilizado foi o modelo de propagação em espaço livre. Uma sessão de dados foi criada entre os nós do canto superior esquerdo [nó 0] e inferior direito [nó 99] utilizando tráfego de taxa de bits constante (CBR) que gera um pacote que é retransmitido pelo receptor para o remetente.

Para testar o movimento, algumas posições de nós foram ajustados para lugares fora de alcance, os nós de 4D a 7G, desaparecem a partir do intervalo de todos os outros nós para testar se o protocolo se adapta à mudança na topologia. É possível observar na Figura 2, que o nó remetente (nó 0) escolhe vizinho mais próximo ao destino. Os vizinhos são os nós na primeira linha, segunda coluna (nó 1) e o nó na segunda linha, primeira coluna (nó 10). Esses dois nós estão à mesma distância do destino e assim o que aparece primeiro na tabela de vizinhos será escolhido, uma vez que a tabela é solicitada pelo IDs dos nós, o nó à esquerda do nó remetente foi escolhido. Em casos onde o nó possui três vizinhos, o nó que está mais próximo do destino é o nó que está abaixo.



Figura 2: Escolha de vizinhos para próximo salto na grade de simulação – Nó 0 escolhe nó 1 ; e nó 1 escolhe nó 11 (O autor, 2014).

Conforme a Figura 3, antes das mudanças de topologia, a comunicação passa pela linha tracejada verde. Após a mudança topológica, as comunicações passam pela arco vermelho. A área sombreada representa os nós que ficam fora de alcance.



**Figura 3: Teste de mobilidade na grade de simulação (Fonseca, Camões, & Vazão, 2012).**

Foram determinados alguns parâmetros para a simulação:

- 30 segundos de tempo de simulação;
- Tamanho do pacote de 1024b;
- Tempo entre envios de pacotes de 1s;

No cenário sem movimento foram obtidos 28 pacotes entregues, com movimento 25 pacotes uma vez que os nós que mudam de posição, movem-se com velocidades de tal forma que os pacotes que estavam com os nós quando se mudaram foram perdidos / descartados. O número 28 é devido que o remetente inicia enviar os pacotes após o momento de 2 s. Ambos os cenários utilizaram 18 saltos para chegar ao destino assim como esperado, tendo em vista o Algoritmo 1.

## 5. Resultados e Considerações

Até o momento o módulo GPSR para o NS-3 ainda não está disponível na versão oficial do simulador, sendo necessário fazer o download do repositório disponibilizado por (Fonseca, Vazão, & Camões, 2011). Devido a simulação ocorrer em um módulo não oficial, foram necessários alguns ajustes para o funcionamento da mesmo. Os testes foram realizados no sistema operacional Ubuntu 9.10 que não possui mais suporte de seu fabricante e nem repositórios oficiais de pacotes, contudo somente essa versão demonstrou compatibilidade com versões das bibliotecas Tcl, Otcl e Tclcl que são essenciais ao módulo GPSR desenvolvido.

Como citado anteriormente, ao contrário do que acontece em protocolos de encaminhamento baseados em topologia, os protocolos de encaminhamento baseados em posição não precisam manter nenhuma rota. A rota é determinada apenas quando for necessário enviar o pacote. Uma outra vantagem é o fato dos protocolos de encaminhamento baseados na posição conterem informações sobre a origem, o destino e

os vizinhos. Sendo assim, como se pode concluir, este tipo de encaminhamento é mais apropriado para VANETs.

Como trabalhos futuros pretende-se realizar um comparativo com outros protocolos de roteamento geográfico para avaliar qual obterá melhor desempenho dentro de cenários de mobilidade em redes VANET.

## Referências

- Coelho, L. (2011). **Avaliação por simulação da comunicação entre veículos**. Universidade de Aveiro. Retrieved from <http://ria.ua.pt/handle/10773/7782>
- Fonseca, A., Camões, A., & Vazão, T. (2012). **Geographical routing implementation in NS3**. Proceedings of the Fifth International Conference on Simulation Tools and Techniques. doi:10.4108/icst.simutools.2012.247688
- Fonseca, A., Vazão, T., & Camões, A. (2011). **Implementation of GPSR on NS3**. Retrieved from <https://code.google.com/p/ns3-gpsr/>
- Jain, R., Puri, A., & Sengupta, R. (2001). **Geographical routing using partial information for wireless ad hoc networks**. Personal Communications, IEEE, 1–25. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=904899](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=904899)
- Karp, B., & Kung, H. (2000). **GPSR: Greedy perimeter stateless routing for wireless networks**. Conference on Mobile Computing and Networking. Retrieved from <http://dl.acm.org/citation.cfm?id=345953>
- Khan, I., & Qayyum, A. (2009). **Performance evaluation of AODV and OLSR in highly fading vehicular ad hoc network environments**. Conference, 2009. INMIC 2009. IEEE 13th ..., 1–5. doi:10.1109/INMIC.2009.5383121
- Li, F., & Wang, Y. (2007). **Routing in vehicular ad hoc networks: A survey**. Vehicular Technology Magazine, IEEE, (June), 12–22. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4450627](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4450627)
- Nzouonta, J., Rajgure, N., & Borcea, C. (2009). **VANET Routing on City Roads Using Real-Time Vehicular Traffic Information**. IEEE Transactions on Vehicular Technology, 58(7), 3609–3626. doi:10.1109/TVT.2009.2014455
- Yousefi, S., Mousavi, M., & Fathy, M. (2006). **Vehicular Ad Hoc Networks (VANETs): Challenges and Perspectives**. 6th International Conference on ITS Telecommunications Proceedings Vehicular. Retrieved from <http://medcontent.metapress.com/index/A65RM03P4874243N.pdf>

# Automatização dos Algoritmos para Estimativa do Balanço de Energia Superficial S-SEBI e SSEB

Fábio Ricardo A. da Silva<sup>1</sup>, Paulo Vinícius S. Benedito<sup>1</sup>, Raphael de Souza R. Gomes<sup>1</sup>, Josiel Maimone Figueiredo<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)  
Cuiabá – MT – Brasil

fabio-ricardo@outlook.com, vini\_programador@hotmail.com,  
{raphael,josiel}@ic.ufmt.br

**Abstract.** *This paper describes the automation process of manually analyzed values in surface energy balance estimative algorithms S-SEBI and SSEB, comparing the results with those obtained by the exhausting method.*

**Resumo.** *Este artigo tem como objetivo descrever o processo de automatização do cálculos de valores analisados manualmente nos algoritmos para estimativa do balanço de energia superficial S-SEBI e SSEB, comparando os resultados obtidos com os do método por exaustão.*

## 1. Considerações Gerais

Recentemente, muitas técnicas afim de estimar a evapotranspiração (*ET*) através do sensoriamento remoto por satélite tem sido desenvolvidas, a exemplo do S-SEBI (*Simplified Surface Energy Balance Index*), proposto por Roerink *et al.* (2000) e do SSEB (*Simplified Surface Energy Balance*), proposto por Senay *et al.* (2011a).

Esses algoritmos proporcionam ao usuário maior praticidade e agilidade, uma vez que podem estimar a *ET* apenas com parâmetro básicos de entrada - como Umidade Relativa e Dia Sequencial do Ano por exemplo - e a imagem de satélite da região desejada, que podem ser obtidos facilmente junto à estação meteorológica e na internet - através dos sites do INPE<sup>1</sup> e do satélite LANDSAT<sup>2</sup>, onde estão disponíveis inúmeras imagens gratuitamente.

Validados em várias partes do mundo, os algoritmos se utilizam dos valores do saldo de radiação (*Rn*), fluxo de calor latente (*LE*), fluxo de calor no solo (*G*) e fluxo de calor sensível (*H*), que, por sua vez, são estimados por meio da aplicação de sucessivas fórmulas sob cada pixel da imagem de satélite da região desejada (ROERINK *et al.*, 2000; SANTOS e SILVA, 2008; SANTOS e SILVA, 2010).

Em contrapartida, o cálculo de algumas variáveis nesses algoritmos se dão a partir de análises manuais feitas por especialistas, tornando difícil o processo de automatização dos mesmos, que é o objetivo deste trabalho.

## 2. Material

Para os testes realizados, utilizou-se imagens do LANDSAT 5-TM adquiridas por meio

- 
- 1 Site oficial do INPE (Instituto Nacional de Pesquisa e Estatística): <<http://www.inpe.br/>>.
  - 2 Site oficial do satélite LANDSAT : <<http://landsat.org/>>.

do site oficial do satélite. Para esse artigo, especificamente, foi utilizada a imagem de órbita 226, ponto 72 do dia 29/09/2011, região do interior de Mato Grosso – Brasil. As implementações foram feitas utilizando a linguagem de programação Python e suas bibliotecas NumPy (OLIPHANT, 2006) e GDAL (*Geospatial Data Abstraction Library*) (AGHISLA, 2014). A máquina usada para os testes possui 4GB de memória, processador Intel® Core™2 Quad CPU Q8200 @ 2.33GHz × 4 e sistema 64 bits.

### 3. Metodologia e Resultados

#### 3.1. Automatização S-SEBI

No S-SEBI, uma das variáveis cujo cálculo é mais importantes para esse algoritmo é a fração evaporativa ( $\Lambda$ ), que é dada pela seguinte fórmula, segundo Gentine *et al.* (2007):

$$\Lambda = \frac{LE}{R_n - G} = \frac{LE}{LE + H} \quad (01)$$

Entretanto, para o cálculo de  $LE$  e  $H$  é necessário ter o valor de  $\Lambda$  previamente. Assim, a fração evaporativa pode ser calculada da seguinte forma (para cada pixel):

$$\Lambda = \frac{T_H - T_0}{T_H - T_{LE}} \quad (02)$$

onde  $T_0$  é a temperatura da superfície (em Kelvin),  $T_H$  e  $T_{LE}$  são as temperaturas correspondentes aos contrastes de seca e umidade para um certo valor de albedo de superfície – valor que indica a reflectância da superfície, variando de 0,0 (escuro) a 1 (brilhante). Aplica-se então a regressão linear em pontos na região das retas  $H_{máx}$  e  $LE_{máx}$  para se obter as equações das retas que representam os limites superiores (máximos) e inferiores (mínimos) (**Figura 1**).

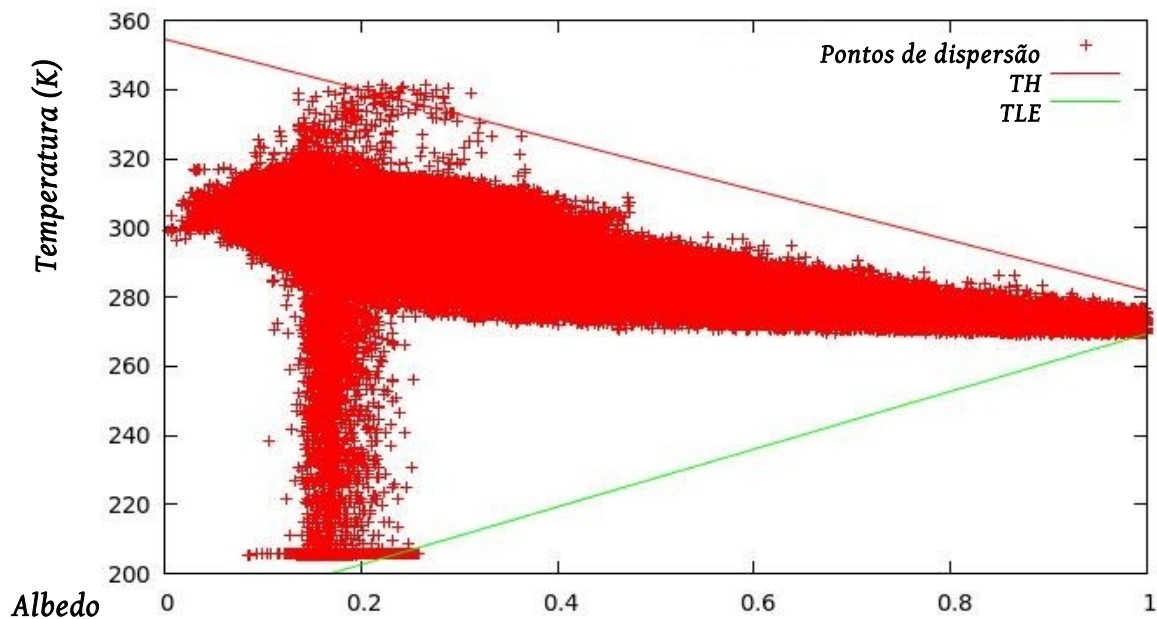


Figura 1- Representação esquemática do gráfico de dispersão de temperatura e albedo de superfície e as retas limite (Método manual).

A construção da **Figura 1** se deu plotando os valores da temperatura pelo albedo superfície, obtidos pela aplicação das fórmulas propostas por Roerink *et al.* (2000) à imagem de satélite. Em seguida, foi realizada, de forma manual, a regressão linear nas extremidades superior e inferior afim de encontrar duas retas de tal forma que formem um tipo de triângulo, evidenciando assim a relação entre as duas variáveis.

Entretanto, o cálculo manual desses valores demanda tempo e cuidado, já que a análise é muito subjetiva, daí a necessidade de automatizar esse cálculo. O método encontrado para tal foi:

1. Considera-se, na **Figura 1**, os pontos nos primeiros 20% dos valores do albedo de superfície (faixa A) e os pontos dos últimos 20% do mesmo eixo (faixa B);
2. De cada faixa, pega-se os 20 (vinte) maiores e os 20 (vinte) menores valores de temperatura e calcula-se a média aritmética para cada grupo de temperatura, resultando em 4 (quatro) valores. Chamaremos de  $y_1, y_3$  as temperaturas alta e baixa, respectivamente, da faixa A e  $y_2, y_4$  as temperaturas alta e baixa, respectivamente, da faixa B;
3. Nessa altura, já se tem todos os valores das ordenadas dos pontos, resta agora os valores das abscissas. Por convenção, para a faixa A, considera-se  $x_1=0,1$  ; para a faixa B, considera-se  $x_2$  como sendo o maior valor do albedo na imagem;
4. Tem-se então os pares ordenados  $(x_1, y_1)$  e  $(x_2, y_2)$  – que serão usados para calcular a reta limite superior – e  $(x_1, y_3)$  e  $(x_2, y_4)$  – que serão usados para calcular a reta limite inferior.

Tendo os quatro pontos, basta agora aplicá-los à definição da equação da reta:

$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0 \Rightarrow x(y_1 - y_2) + y(x_2 - x_1) + (x_1 y_2 - x_2 y_1)$$

Fazendo  $(y_1 - y_2) = a$  ,  $(x_2 - x_1) = b$  e  $(x_1 y_2 - x_2 y_1) = c$  , como  $a$  e  $b$  não são simultaneamente nulos , temos:

$$ax + by + c = 0$$

(equação geral da reta)

Isolando  $y$  na equação geral, teremos a equação reduzida da reta:

$$y = \frac{-a}{b}x - \frac{c}{b}$$

$$m = \frac{-a}{b}$$

$$q = \frac{-c}{b}$$

Voltando ao S-SEBI, as equações reduzidas das retas limite serão dadas por:

$$T_H = a_H + b_H \alpha_0 \quad (03)$$

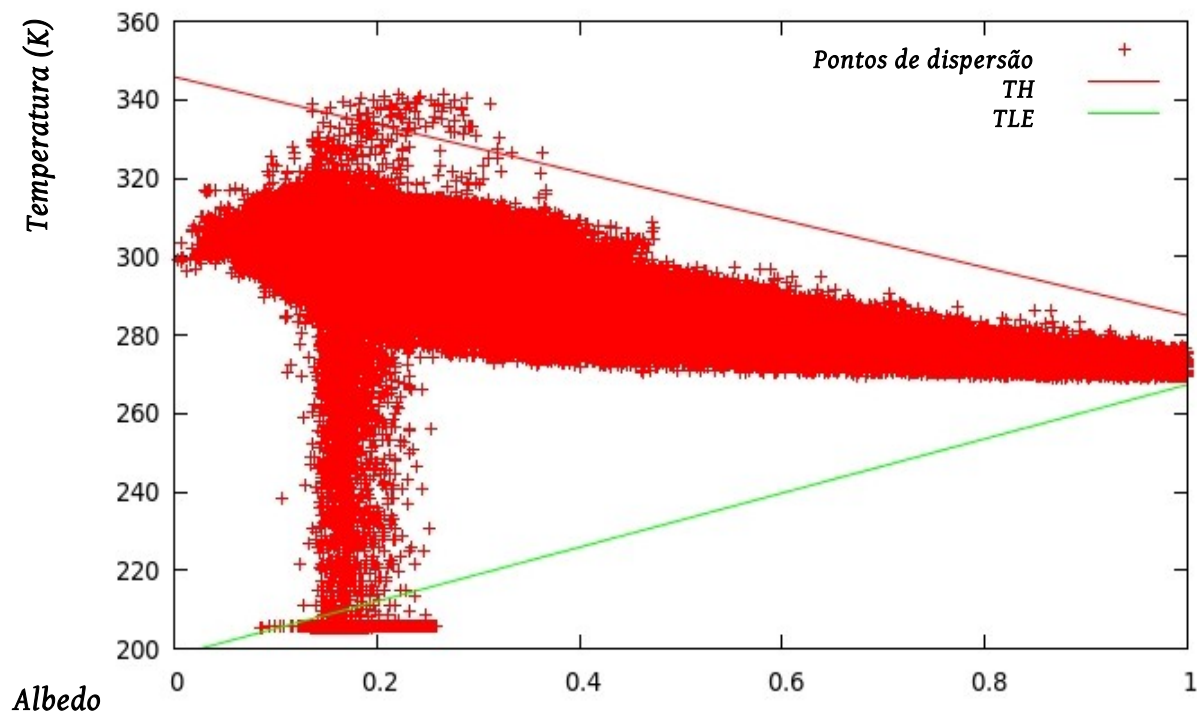
$$T_{LE} = a_{LE} + b_{LE} \alpha_0 \quad (04)$$

onde  $a_H$ ,  $a_{LE}$ ,  $b_H$  e  $b_{LE}$  são os coeficientes das retas encontradas e  $\alpha_0$  é o albedo da superfície.

Substituindo as equações (03) e (04) em (02), tem-se:

$$\Lambda = \frac{a_H + b_H \alpha_0 - T_0}{a_H - a_{LE} + (b_H - b_{LE}) \alpha_0} \quad (05)$$

Após os cálculos, obteve-se o seguinte gráfico( **Figura 2**):



**Figura 2- Representação esquemática do gráfico de dispersão de temperatura e albedo da superfície e as retas limite (Método automatizado).**

### 3.2. Automatização SSEB

Já no algoritmo SSEB, a variável de cálculo manual é a fração de evapotranspiração ( $ET_f$ ), que utiliza o artifício de encontrar a média dos 3 (três) pixels-âncora quentes ( $T_H$ ) e a média 3 (três) pixels-âncora frios ( $T_C$ ) de acordo com os valores do NDVI (*Normalized Difference Vegetation Index*), ou seja, para um dado período, o pixel frio representa culturas bem vegetadas e bem regadas, selecionado pela

combinação de valores baixos de  $T_0$  e altos valores de NDVI. Já o pixel quente representa vegetação de baixa densidade e solo relativamente seco, identificado por altos valores de  $T_0$  e baixos valores de NDVI (SENAY *et al.*, 2007; 2011b).

O cálculo da fração de evapotranspiração é dado pela seguinte expressão:

$$ET_f = \frac{(T_H - T_0)}{(T_H - T_C)} \quad (06)$$

onde  $T_H$  é a média aritmética dos três pixels-âncora quentes,  $T_C$  é a média aritmética dos três pixels-âncora frios e  $T_0$  a temperatura de cada pixel.

Porém, essa seleção manual dos pixels torna o processo lento e trabalhoso. Logo, o método encontrado para automatizar esse cálculo foi:

1. Armazena-se, da imagem do NDVI, as posições correspondentes aos três maiores valores de NDVI (para o pixel frio) e aos três menores valores de NDVI (para o pixel quente);
2. Feito isso, encontra-se, em  $T_0$ , as temperaturas correspondentes às posições encontradas anteriormente;
3. Tendo-as em mãos, calcula-se então a média aritmética das temperaturas cujo NDVI foi alto ( $T_C$ ) e a média das temperaturas cujo NDVI foi baixo ( $T_H$ ).

#### 4. Conclusão

Utilizando uma amostra de 5 (cinco) imagens, os resultados obtidos através da automatização desses cálculos nos algoritmos S-SEBI e SSEB foram satisfatórios. No caso do S-SEBI, a estimativa dos coeficientes das retas pelo método automatizado atingiu até 85,4% de precisão em relação ao método manual nas imagens testadas.

Já no SSEB, a precisão foi de 100%, já que o algoritmo – ao contrário de seu similar S-SEBI – não trabalha com um cálculo propriamente dito, mas com uma seleção de pixels, por isso consegue resultados idênticos aos obtidos através do método manual, ou seja, o computador consegue identificar exatamente os extremos da temperatura.

Portanto, as automatizações no S-SEBI e SSEB para a obtenção da fração evaporativa e pixels-âncora, respectivamente, mostraram-se ser boas alternativas de cômputo dos mesmos, uma vez que retira a necessidade de análises manuais e as automatiza. Desta forma, contribui para evitar erros e tendências dos analistas.

## Referências Bibliográficas

- AGHISLA, (2014) “GDAL/OGR 1.11.0 released”, <<http://www.osgeo.org/node/1442>>.
- GENTINE, P., ENTEKHABI, D., CHEHBOUNI, A., BOULET, G., DUCHEMIN, B. (2007) “Analysis of evaporative fraction diurnal behavior.”, *Agricultural and Forest Meteorology*, v.143, p. 13 – 29.
- GOWDA, P.H., SENAY, G. B., HOWELL T. A. and MAREK, T. H. (2009) “Lysimetric Evaluation of Simplified Surface Energy Balance Approach in the Texas High Plains.”, v.25, p 665-669.
- OLIPHANT, T. E. (2006) “Guide to NumPy”, <<http://csc.ucdavis.edu/~chaos/courses/nlp/Software/NumPyBook.pdf>>, Dezembro.
- ROERINK, G. J., SU, Z and MENENTI, M. (2000) “S -SEBI: a simple remote sensing algorithm to estimate the surface energy balance.”, *Physics and Chemistry of the Earth (B)*. n. 25, p. 147-157.
- SANTOS, C. A. C., SILVA, B. B. (2008) “Estimativa da Evapotranspiração da Bananeira em Região Semiárida Através do Algoritmo S-SEB.”, *Revista Brasileira de Agrometeorologia*, v.16, p. 9 – 20.
- SANTOS, C. A. C., SILVA, B. B. (2010) “Obtenção dos fluxos de energia à superfície utilizando o algoritmo S-SEBI.”, *Revista Brasileira de Agrometeorologia*, v.25, p. 365 – 374.
- SENAY, G. B., BUDEE, M. E., VERDIN, J. P. and MELESSE, A. M. (2007) “A Coupled Remote Sensing and Simplified Surface Energy Balance Approach to Estimate Actual Evapotranspiration from Irrigated Fields.”, *Sensors*, v. 7, p. 979-1000.
- SENAY, G. B., BUDEE, M and VERDIN, J. P. (2011a) “Enhancing the Simplified Surface Energy Balance (SSEB) approach for estimating landscape ET: Validation with the METRIC model.”, *Agricultural Water Management*, v. 98, p. 606-618.
- SENAY, G. B., LEAKE, S., NAGLER, P. L., ARTAN, G., DICKINSON, J., CORDOVA, J. T. and GLENN, E. P. (2011b) “Estimating basin scale evapotranspiration (ET) by waterbalance and remote sensing methods.”, *Hydrological Processes*, v. 25, p. 4037-4049.

# Identificação de Eventos de Áudio para Monitoramento de Aves no Pantanal Mato-grossense

Alan Rodrigo S. Marcon<sup>1</sup>, Nielsen Cassiano Simões<sup>1</sup>, Josiel Maiomone de Figueiredo<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)  
Rua Fernando Correa da Costa, 2367 - Boa Esperança – Cuiabá – MT – Brasil

alanrodrigo@ic.ufmt.com.br, nielsen@ic.ufmt.br, josiel@ic.ufmt.br

**Abstract.** *The sound analysis of animals in the pantanal can tell us several aspects about their behavior, studying them will provide useful information in the scientific area. However with a rich fauna and 24 hours of recording, manual analysis of these data is not feasible, at this point the computational detection of events in an audio file make it easy the process of analysis. This paper presents a method to perform this detection and separation of events from data obtained in the Patanal area of the Mato Grosso State.*

**Resumo.** *A análise dos sons produzidos pelos animais do Pantanal pode nos dizer muito sobre seu comportamento, estudá-los irá prover-nos muitas informações úteis na área científica, mas com uma fauna tão rica, a análise manual desses dados é inviável, neste ponto a detecção computacional de eventos em um arquivo áudio agiliza em muito o processo de identificação. Este artigo apresenta um método para realizar esta detecção e a separação de eventos a partir de dados captados no Pantanal mato-grossense.*

## 1.Introdução

A fauna do Pantanal é privilegiada contando com inúmeras espécies tanto endêmicas quanto migratórias, destacando-se neste último grupo as aves. Conhecer e entender seu comportamento é de suma importância para avanços científicos nas áreas biológicas.

Diversos projetos relacionados ao INAU (Instituto Nacional de Áreas Úmidas) possuem a tarefa de identificar e guardar dados a respeito da fauna pantaneira, em uma área de aproximadamente 150.355 km<sup>2</sup>, habitada por inúmeras espécies, das quais cerca de 463 são aves (Ambiente, 2014).

Neste trabalho são apresentados alguns aspectos relacionados a resultados obtidos pelo Grupo de Pesquisa “*Computational Bioacoustic Resource Unit*” (CO.BRA), cujo objetivo principal é o monitoramento da biodiversidade do Pantanal a partir de dados de bioacústica, o que neste artigo relaciona-se com o desenvolvimento de ferramentas computacionais automatizadas para esse fim<sup>1</sup>.

Na região de Pouso Alegre próximo a Poconé/MT, foram implantadas oito bases para a captação de sinais sonoros deste bioma, espalhadas em pontos estratégicos, compostas por dois

---

<sup>1</sup> <http://cobra.ic.ufmt.br>

microfones de alta definição que efetuam o registro contínuo 24/24 horas. O volume de dados gerados é de aproximadamente 16GB por base semanalmente.

A análise manual dos dados é eficaz na separação de eventos julgados importantes, porém a dimensão deste volume de dados provoca uma ineficiência deste processo, principalmente, em virtude do tempo gasto para a análise e fadiga do operador.

Buscando otimizar esta fase do processo, podemos utilizar algoritmos que identificam possíveis eventos em um trecho de áudio, visando diminuir a quantidade absoluta de dados que serão analisados para uma posterior identificação das espécies.

## 2. Ferramentas

Todo o processo é feito utilizando a linguagem Python em conjunto com ferramentas e bibliotecas de terceiros, sendo estas: *Modal (Musical Onset Database And Library)* (GLOVER et al. 2011) utilizada para a identificação de *onsets*; *scipy* utilizada para tratar dados científicos; e *numpy* (JONES Eric et al. 2001) responsável por cálculos matemáticos, principalmente aqueles que envolvem matrizes, no caso o arquivo *wav* que é interpretado pelo computador como uma matriz.

## 3. Processo de identificação

Os dados obtidos pelas bases de captura de áudio no Pantanal estão armazenadas no formato WAV (*Wave Audio File Format*), compostos por dois canais (estéreo), em 48KHz e 32bits, com duração aproximada de 14 minutos por arquivo.

O processo de análise dos dados divide-se em três etapas: pré-processamento, identificação de eventos e separação dos eventos detectados, como podem ser observadas na Figura 1.

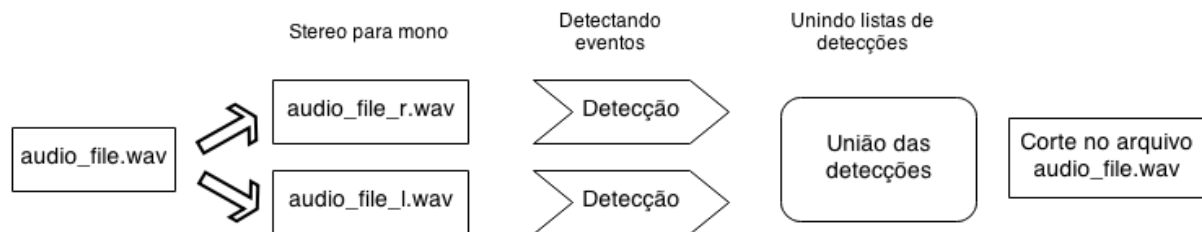


Figura 1: Etapas no processo de identificação de eventos

### 3.1. Pré-processamento

O pré-processamento envolve a análise do nome do arquivo que contém informações sobre a base de captação, data e horário das gravações. Nesta etapa é realizada uma filtragem de passa alta e passa baixa e a separação dos canais de áudio do arquivo *wav* – canal esquerdo e canal direito – para que o arquivo possa ser analisado pelos outros algoritmos.

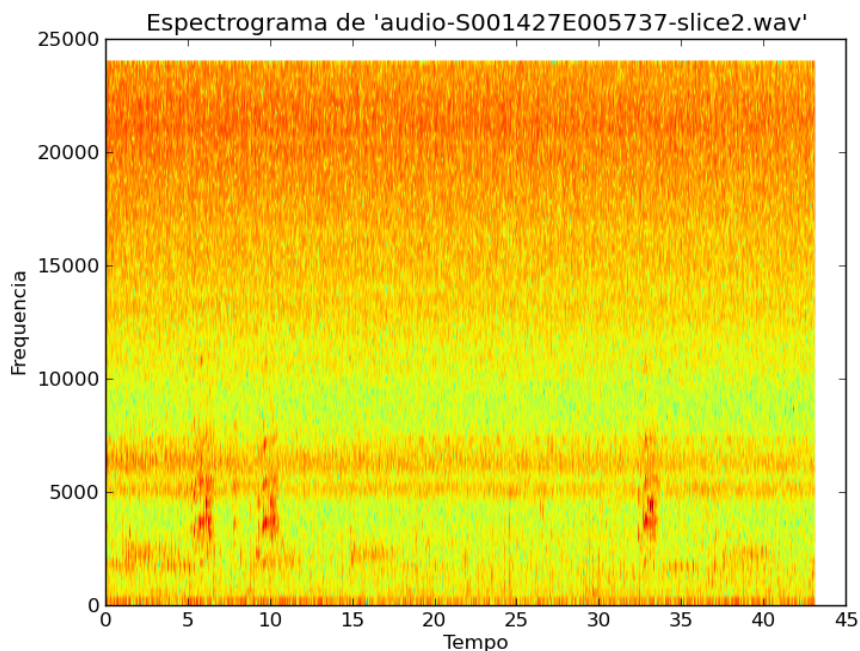
A filtragem consiste na eliminação de determinadas frequências, nas quais não teremos eventos úteis por conterem um alto nível de ruído (normalmente gerado por ventos) ou por serem frequências as quais o ouvido humano não irá captar. Esta etapa do processo é opcional, sua necessidade deve ser verificada conforme o conjunto de dados analisados. Por padrão são removidas as frequências abaixo de 300KHz e aquelas acima de 20000KHz.

Inicialmente para executar a extração de canais, eram utilizadas funções das bibliotecas nativas da linguagem Python. Entretanto, este método demandava muito tempo de processamento, o qual, para arquivos de áudio relativamente pequenos (com cerca de 4 a 10 MB), é pouco relevante, porém, como o arquivo de trabalho que é utilizado possui cerca de 161,3 ou 345,1 MB, combinado com o volume de arquivos a ser processado, torna todo o processo custoso.

Buscando melhorar este desempenho, após alguns estudos, passamos a utilizar a biblioteca *scipy* (JONES Eric et al. 2001) para o tratamento direto dos dados, conseguindo executar a mesma tarefa, mas, com melhor desempenho, como será observado no capítulo de resultados.

### 3.2. Detecção de eventos

Após o pré-processamento, é necessário que os eventos sejam detectados para que trechos de silêncio (segmentos de baixa energia) possam ser eliminados, buscando reduzir o conjunto de dados. A Figura 2 ilustra o espectrograma de um arquivo. Note que os eventos de interesse possuem uma coloração de destaque.



**Figura 2: Espectrograma gerado em python contendo vários eventos**

A detecção consiste em um conjunto de diferentes códigos ODF (*Onset Detection Function* – detecção do início de eventos). O processo se realiza da seguinte forma:

1- Leitura do arquivo de áudio.

```
arquivo_de_trabalho = arquivo_audio_lido
```

```
audio, info = características(arquivo_de_trabalho)
audio = audio / info.valormáximo
```

2- Definição do tamanho da janela de análise e o tamanho de cada passo.

```
tamanho_do_frame = 960
tamanho_do_passo = 340
```

3- Computar as detecções e gravá-las em uma lista.

```
enquanto existe_frames_a_processar == verdadeiro:
    analisa_frames
    realiza_detecção
    se houve_detecção então:
        armazena_frame_na_lista
    da_mais_um_passo
```

Podemos demonstrar graficamente onde as detecções foram encontradas após este processo, como pode ser observado na Figura 3, na qual cada linha tracejada vermelha indica a localização do *frame* em que ocorreu uma detecção.

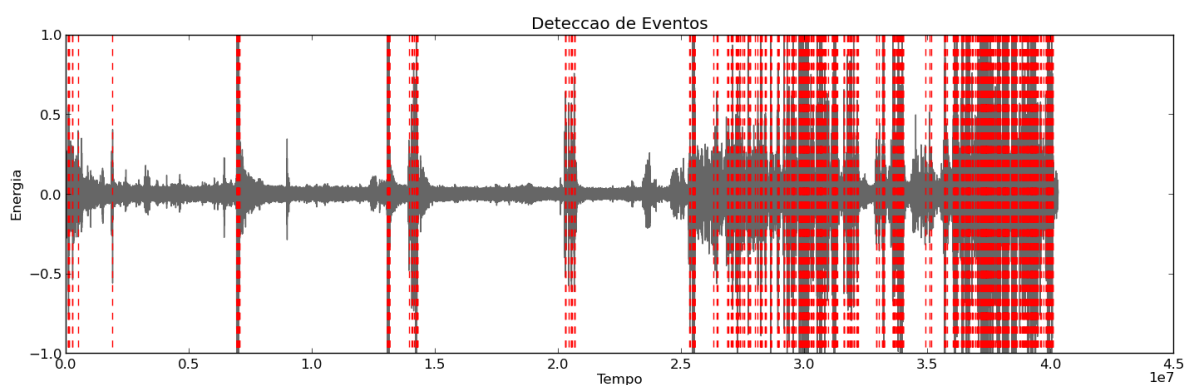


Figura 3: Exemplo 1 de plote das detecções

### 3.3. Junção e corte

Após obtermos uma lista com a indicação dos *frames* das detecções, devemos estabelecer o tamanho dos cortes que serão realizados no áudio de entrada. A Figura 4 ilustra este processo.

Como o código trabalha com arquivos mono, devemos processar duas faixas de áudio para cada arquivo de entrada, buscando obter o máximo de fidelidade possível aos possíveis eventos captados, feito isto obtemos duas listas de detecções. Para cada lista, estabelecemos o que chamamos de *background*, que é um intervalo de tempo antes e após cada detecção. Um *background* pode englobar outra detecção, provocando uma junção de ambas. Portanto o novo ponto de fim do *background* será recalculado. As detecções provenientes de um mesmo arquivo

apresentam diferenças, pois os microfones que os captam estão posicionados em direções opostas. Dessa forma, criamos uma nova lista resultante da combinação das duas listas iniciais já com o *background*.

Ao obtermos esta lista geral de detecções, realizamos os cortes dos segmentos resultantes, tendo ao final vários arquivos de áudio estéreo ou mono no formato *wav*, conforme a necessidade do operador.

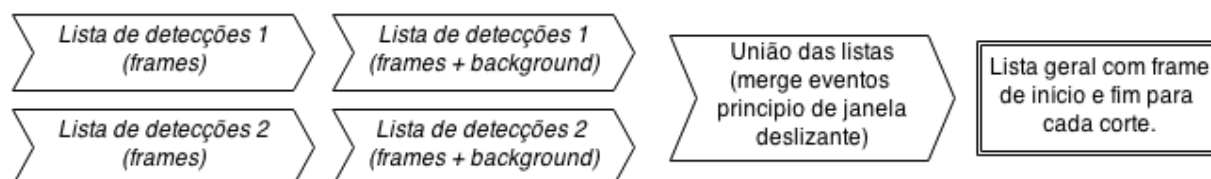


Figura 4: Processo de criação das listas de detecções

#### 4.Resultados

O computador utilizado neste projeto possui as seguintes configurações: processador Intel® Core™2 Quad CPU Q8200; 4GB de memória RAM DDR 3 800MHz, HD 240GB SATA II 5400rpm.

Na etapa de pré-processamento, o método inicial realizava a extração dos canais em aproximadamente 13 minutos. O novo método utilizado reduziu este tempo para algo em torno de 5 segundos (99,36% inferior) isso dá-se principalmente ao fato de que a biblioteca *scipy* através de suas funções de leitura e escrita de arquivos *wav*, interpreta cada canal como a coluna de uma matriz, e no momento da gravação em disco é feita a gravação da coluna da matriz, diretamente em um arquivo independente, dispensando uma análise dos *frames* e características como era feito pelo método inicial.

As concepções de um evento são estabelecidas por pesquisadores associados ao CO.BRA, e para todos os casos de teste aqueles eventos considerados relevantes por eles foram detectados, porém com variações em relação ao *background*, pois na detecção manual existe o fator visual do espectrograma que pesa no momento da separação.

A Tabela 1 exibe um comparativo do tamanho dos dados, antes e após o processamento, e a taxa de diminuição verificada. Note que, em média a taxa de redução foi de 54,37%.

Tabela 1. Comparativo no tamanho total dos arquivo antes e após a identificação de eventos relevantes

| Nome do arquivo | Tamanho inicial (MB) | Tamanho processado(MB) | Número de cortes | Taxa de diminuição |
|-----------------|----------------------|------------------------|------------------|--------------------|
| teste1.wav      | 161,3                | 20,9                   | 10               | 87,04%             |
| teste2.wav      | 161,3                | 133,6                  | 22               | 17,17%             |
| teste3.wav      | 161,3                | 34,1                   | 15               | 78,86%             |
| teste4.wav      | 161,3                | 29,8                   | 12               | 81,52%             |
| teste5.wav      | 161,3                | 45,2                   | 11               | 71,97%             |
| teste6.wav      | 161,3                | 58,7                   | 18               | 63,60%             |

|               |               |              |            |               |
|---------------|---------------|--------------|------------|---------------|
| teste7.wav    | 161,3         | 111,8        | 18         | 30,68%        |
| teste8.wav    | 161,3         | 154,7        | 30         | 4,09%         |
| <b>Totais</b> | <b>1290,4</b> | <b>588,8</b> | <b>136</b> | <b>54,37%</b> |

Percebe-se que para alguns arquivos a diminuição é acima de 80% já em outros abaixo de 5% isso ocorre devido ao horário no qual as gravações foram feitas, pois existem momentos durante o dia/noite com baixos níveis de atividade sonora devido aos períodos de atividade da fauna, bem como momentos com altos níveis de captação, seja pelo grande número de animais dentro da área de alcance dos microfones, ou por adventos naturais como chuva ou vento.

## 5. Conclusão

A etapa de detecção de eventos de áudio, captados na natureza, é dificultada por não haver uma filtragem prévia do que irá ser analisado, ou seja, obtemos áudios em seus formatos brutos, com altos níveis de ruído, completamente suscetíveis a adventos naturais, tais como chuva e vento. Além disso, é comum ocorrerem eventos coincidentes, ou seja, mais de um animal tem seu som captado ao mesmo tempo, dificultando o processo de detecção dos eventos. Logo uma intersecção de detecções é interpretado como um evento único. Com a aplicação destas técnicas, o tamanho da quantidade de dados para análise será menor, o que agiliza processamentos e análises futuras destes dados.

O desenvolvimento deste trabalho tornou-se possível graças ao apoio do Instituto de Computação da UFMT campus Cuiabá, em parceria com o INAU (Instituto Nacional de Áreas Úmidas) tendo o apoio dos Grupos de Pesquisa “*Computational Bioacoustic Resource Unit*” (CO.BRA), e GAIIA (Gerenciamento e Armazenamento Inteligente de Imagens Ambientais).

## Referências

- AMBIENTE, Ministério do Meio. (2014) Pantanal, <<http://www.mma.gov.br/biomas/pantanal>>, acessado em agosto 2014.
- GLOVER, John ; LAZZARINI, Victor; TIMONEY, Joseph. et al. (2011) Real-Time Detection of Musical Onsets with Linear Prediction and Sinusoidal Modelling, EURASIP Journal on Advances in Signal Processing, 2011:68.
- JONES Eric, OLIPHANT, T. E., PETERSON Pearu, et al. (2001). SciPy Open source scientific tools for Python. <<http://www.scipy.org>>

## Comparação do desempenho entre banco de dados não relacional e relacional em análise de patentes

Phellippe Lorrann Pinheiro Lima<sup>1</sup>, Cristhian de Melo Ferreira<sup>1</sup>, Mattyws Ferreira Grawe<sup>1</sup>, Raphael de Souza Rosa Gomes<sup>1</sup>, Josiel Maimone de Figueiredo<sup>1</sup>, Andreia Bonfante<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)  
Cuiabá – MT – Brasil

{phellippe.end, cmferreira.352, mattyws}@gmail.com,  
{raphael, josiel, andreia.bonfante}@ic.ufmt.br

**Abstract.** *This paper's objective is to compare the query execution performance between two database types, the relational and the non relational (NoSQL), when used for patent's information storage. As well as the explanations about what are non relational databases and the reasons they were used for the system implementation.*

**Resumo.** *Este artigo tem o objetivo de comparar o desempenho na execução de consultas entre dois tipos de banco de dados, o relacional e o não relacional (NoSQL), quando utilizados para o armazenamento de informações de patentes. Assim como explicações sobre o que são bancos de dados não relacionais e por que foram usados para a implementação do sistema.*

### 1. Introdução

O Periscope é um projeto que tem como objetivo principal disponibilizar para a comunidade brasileira uma ferramenta que auxilia o armazenamento e a análise das informações contidas no Sistema de Propriedade Industrial [Viebrantz 2012].

Segundo a Organização Mundial da Propriedade Intelectual (OMPI), cerca de 2/3 das informações tecnológicas disponíveis em todo mundo somente são contidas nos documentos de patentes. Atualmente no mundo o número de documentos de patentes cresce cerca de 1.500.000 por ano [Rodrigues 2012].

Estudo feito na Alemanha concluiu que os custos de P&D (Pesquisa e Desenvolvimento) poderiam ser reduzidos em 30%, caso as informações técnicas disponíveis fossem utilizadas. Segundo estimativa da Organização Britânica de Patentes, 20 bilhões de libras por ano são desperdiçadas na Comunidade Europeia devido a invenções duplicadas.

Isso acontece porque uma patente ao ser registrada leva de 3 a 5 anos para ser publicada efetivamente, no entanto, a partir do momento em que uma patente for publicada, seus dados ficam disponíveis para o acesso público em repositórios de patentes [PatentSCOPE 2014], [DPMA 2014], [Espacenet : EPO 2014].

O sistema oferece ao usuário gráficos com dados para análise das patentes que o mesmo adicionar ao seu projeto, tecnologias como: JAVA EE, JBOSS, NetBeans, Git, Apache Maven e MongoDB foram utilizadas para o desenvolvimento.

O MongoDB é um banco de dados não relacional orientado à documento em JSON com esquemas dinâmicos, *open-source*, escrito em C++, também é o banco de

dados NoSQL mais usado da atualidade como é informado no site oficial<sup>1</sup> do próprio banco.

## 2. Porque NoSQL

Banco de dados NoSQL ou 'Not Only SQL', representam a nova classe de tecnologias de gerenciamento de dados para suprir o aumento no volume, velocidade e variedade dados que organizações estão armazenando, processando e analisando<sup>2</sup>.

Existem duas possíveis razões para mudança para bancos de dados não relacionais: performance e flexibilidade [Stonebraker 2010]. Comparados com bancos de dados relacionais, bancos de dados NoSQL escalam melhor e fornecem performance superior. Bancos de dados NoSQL oferecem as oportunidades que bancos de dados relacionais não oferecem, incluindo:

- Grandes volumes de dados estruturados, semi-estruturados e não estruturados;
- Flexível, fácil de usar com programação orientada a objeto;
- Eficiente, escala horizontalmente ao invés de vertical como os relacionais.

O MongoDB foi escolhido, pois, além de ser o banco de dados não relacional mais usado da atualidade [Mongo 2014], possui algumas das principais funcionalidades necessárias para o desenvolvimento do sistema.

Por trabalhar com grandes quantidades de dados, o Periscope seria um sistema extremamente lento se os dados fossem armazenados por um banco de dados relacional. O banco de dados relacional se torna muito lento por conta de suas tabelas, chaves estrangeiras e *joins*.

A **Figura 1** mostra o Diagrama de Entidade Relacionamento do banco de dados se as informações da patente fossem armazenados em um banco de dados relacional, enquanto a **Figura 2** mostra um exemplo de um documento de dados em JSON onde estão armazenados as mesmas informações das mesmas patentes.

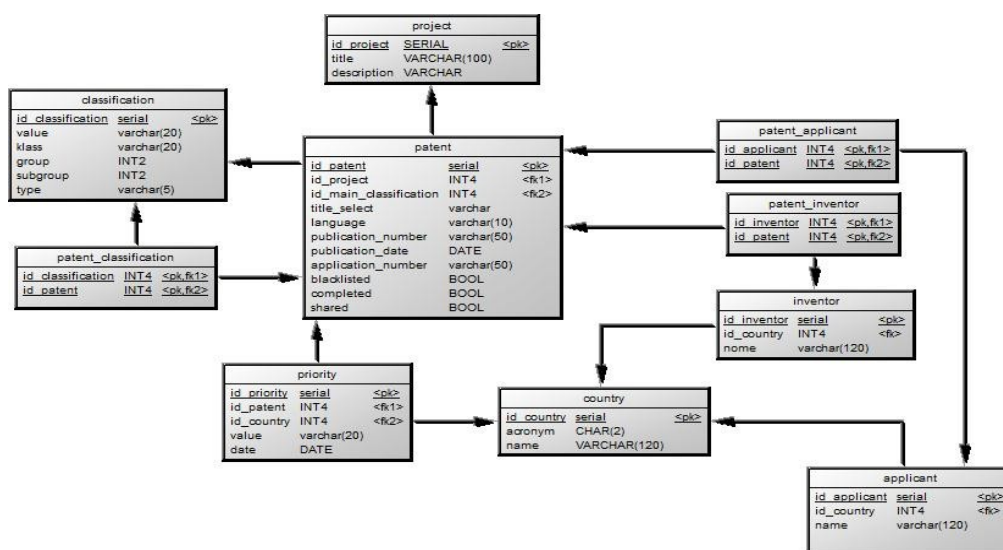


Figura 1. Armazenamento de patentes em banco relacional

- 1 Site Oficial MongoDB <<http://www.mongodb.com/leading-nosql-database>>
- 2 Site MongoDB <<http://www.mongodb.com/nosql-explained>>

```

{
  "_id" : ObjectId("53e0e43644ae03b10fb4b00c"),
  "titleSelect" : "Titulo da patente",
  "mainClassification" : {
    "value" : "Valor", "klass" : "Classe", "group" : "Grupo", "subgroup" : "Subgrupo"
  },
  "classifications" : [
    {Classificação1}, {Classificação2}
  ],
  "priorities" : [
    {"value" : "Valor", "country" : {"acronym" : "PA", "name" : "PAÍS"},
    "date" : Data},
    {prioridade2}
  ],
  "language" : "EN", "publicationNumber" : "Numero", "publicationDate" : Data,
  "applicationNumber" : "Numero", "applicationDate" : Data,
  "applicants" : [
    {"name" : "Depositante", "country" : {"acronym" : "PA", "name" : "PAÍS"}},
    {Depositante2}
  ],
  "inventors" : [
    {"name" : "Inventor", "country" : {"acronym" : "PA", "name" : "PAÍS"}},
    {Inventor2}
  ],
  "blacklisted" : false,
  "completed" : true,
  "shared" : false,
  "project" : DBRef("Project", ObjectId("53e0e40a44ae03b10fb4affb"))
}

```

**Figura 2. Modelo de um JSON armazenando valores de uma patente**

### 3. Testes

O banco de dados relacional escolhido para os testes foi o PostgreSQL, a conexão com o sistema é feita pelo JPA (Java persistence API), contra o não relacional usado no Periscope, MongoDB. Ambos foram populados com os dados de cerca de 100.000 patentes, cada patente possui as informações dos depositantes, inventores, classificações e prioridades.

No Periscope existem classes responsáveis por importar as informações da patente e armazená-las no banco de dados, gráficos são gerados para detalhar melhor essas informações de tal forma que análises possam ser feitas pelos gráficos.

Os principais gráficos são:

1. Teste 1: Principais depositantes, quantidade de patentes que cada depositante está envolvido;
2. Teste 2: Principais inventores, quantidade de patentes que cada inventor está envolvido;
3. Teste 3: Principais classificações, quantidade de patentes por classificação;
4. Teste 4: Principais países de prioridades, quantidade de patentes que cada país tem como prioridade.

A consulta no MongoDB para gerar o primeiro gráfico é a seguinte:

```

db.Patent.aggregate({$unwind:"$applicants"},
  {$group: {_id:"$applicants", applicationPerApplicant: {$sum:1}}},
  {$sort: {applicationPerApplicant:-1}})

```

A consulta no PostgreSQL para gerar o primeiro gráfico é a seguinte:

```
SELECT applicant.name,
       sum(1)
FROM applicant
JOIN patent_applicant USING (id_applicant)
GROUP BY applicant.name
ORDER BY SUM DESC;
```

O tempo de importação das informações das 100.000 patentes para cada banco de dados também foi comparado.

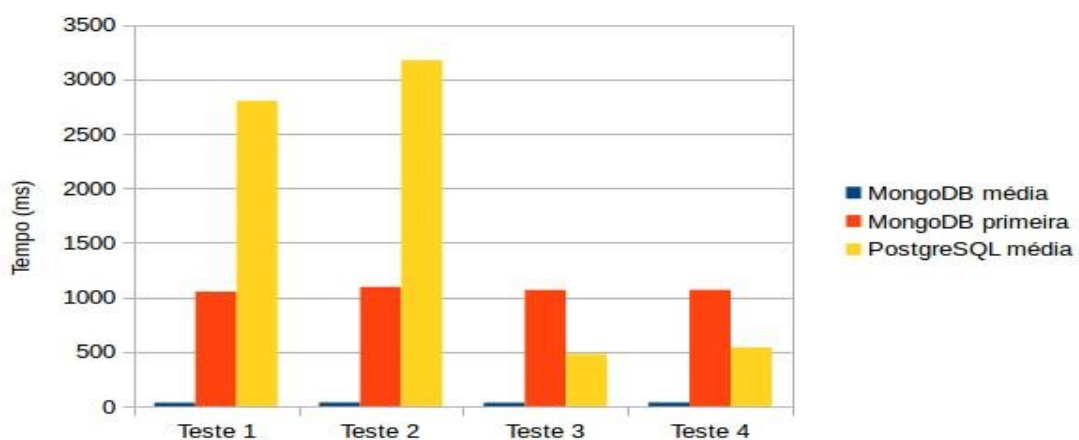
#### 4. Resultados

O computador utilizado neste projeto possui as seguintes configurações: processador Intel® Core™2 Quad CPU Q8200; 4GB de memória RAM DDR 3 800MHz, HD 320GB SATA II 5400rpm. O sistema operacional utilizado foi o Ubuntu 14.04 LTS 64 bits.

A **Figura 3** mostra o resultado do tempo de execução de cada um dos testes para os dois bancos de dados. O MongoDB aparece duas vezes, sendo a primeira consulta a única que realmente leva um tempo maior, todas as consultas subsequentes estão na faixa de 32 à 34 ms.

Isso acontece por que o MongoDB armazena as consultas em cache, fazendo com que todas as consultas depois da primeira se tornem extremamente mais rápidas, não somente aquela consulta específica, mas todas consultas naquele banco que esta sendo usado, sendo assim, no Periscope, só é necessário carregar um gráfico e o tempo de execução de todas as outras consultas vai ser muito pequeno.

Se comparado ao tempo médio de execução dos testes do PostgreSQL, independente da quantidade de vezes uma consulta é executada a média de execução continua elevada.



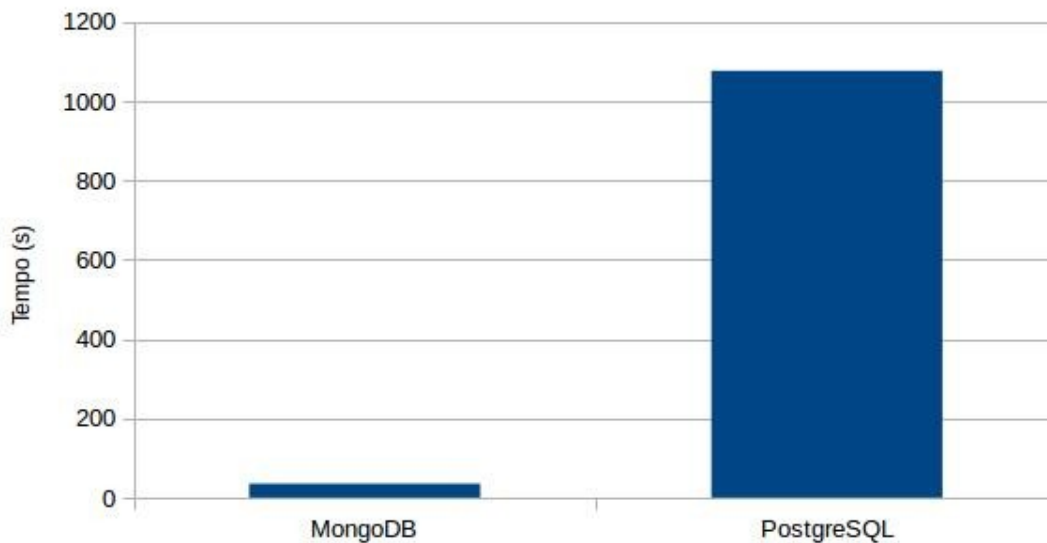
**Figura 3.** Gráfico com os tempos de execução dos testes para cada banco de dados, a descrição de cada teste está na seção anterior.

A **Figura 4** mostra o tempo de importação das 100.000 patentes usando cada banco de dados, a importação em ambos os bancos foi feita um a um, ou seja, uma patente foi inserida no banco de cada vez.

A inserção dos dados no MongoDB é feita colocando todas informações de uma patente em um único documento de uma coleção de patentes, como foi dado o exemplo da **Figura 2**.

No PostgreSQL, no entanto, existem diversas tabelas para cada tipo de dado de uma patente (**Figura 1**), e essas tabelas geram chaves estrangeiras, o que deixa não só o processo de inserção mais lento, mas também as consultas. Além disso antes de adicionar os dados o sistema deve verificar se aquela informação já existe no banco de dados, para evitar a replicação de dados desnecessária, o que não acontece no MongoDB.

As informações que podem estar replicadas no MongoDB, como nome de depositantes, inventores, classificações, são replicadas sem nenhum tratamento prévio, por conta disso, existe um gasto maior de disco para armazenar essas informações, mas os gráficos provam que o benefício disso é a velocidade que as consultas são executadas.



**Figura 4. Tempo de importação de 100.000 patentes nos bancos de dados**

## 5. Conclusão

O MongoDB provou ser superior ao PostgreSQL em todos os testes executados, considerando que, mesmo nas consultas para gerar os gráficos em que aparentemente o MongoDB foi mais lento que o PostgreSQL, isso na verdade só vai acontecer uma única vez, se por acaso o primeiro gráfico que o usuário abrir ao entrar no sistema for um dos gráficos dos Testes 3 e 4, mas, depois que a primeira consulta é executada, o MongoDB mantém uma média de tempo de execução muito superior que o PostgreSQL.

O MongoDB também se provou muito superior ao PostgreSQL no teste de tempo de importação, mesmo com a importação não otimizada, sendo uma patente adicionada ao banco por vez, ao invés de por blocos, com vários patentes sendo

adicionadas de uma só vez, como o sistema foi implementado.

Logo, como o sistema foi criado para auxiliar a análise das informações contidas em patentes, como foi dito no item 1, o tempo de carregamento de cada um dos gráficos deve ser o mínimo possível. O MongoDB por trabalhar com documentos sem a necessidade de *updates*, somente adicionando os dados e depois fazendo buscas, isso mostrou-se ser mais eficiente e portanto a melhor escolha.

Para o futuro, pretende-se realizar testes em casos diferentes, utilizando otimização de consulta em ambos os bancos de dados, e também usando diferentes paradigmas de NoSQL, como orientado a grafos, colunas e outros bancos de dados orientados a documentos, de forma verificar a performance de cada um, bem como sua utilização.

## 6. Agradecimentos

Essa pesquisa foi apoiada em parte pela Universidade Federal de Mato Grosso (UFMT) e pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

## Referências Bibliográficas

- DPMA (2014) “DPMA – Deutsches Patent- und Markenamt” <<http://www.dpma.de/english>> acessado em agosto de 2014.
- Espacenet: EPO (2014) “Espacenet : European Patent Office” <<http://www.epo.org/searching/free/espacenet.html>> acessado em agosto de 2014.
- Mongo (2014) “MongoDB – The Leading NoSQL Database” <<http://www.mongodb.com/leading-nosql-database>> acessado em agosto de 2014.
- PatentSCOPE (2014) “PatentSCOPE – Search International and National Patent Collections” <<http://patentscope.wipo.int/>> acessado em agosto de 2014.
- Rodrigues, R. C. (2012) “INPITEC uma ferramenta gratuita para tratamento estatístico de informações oriundas do sistema de patentes” <<http://pt.slideshare.net/ricardocarvalhorodrigues/inpitec-cuiab>> acessado em agosto de 2014.
- Stonebraker M. (2010) “SQL databases v. NoSQL Databases” , Communications of the ACM, v. 53, p 10 – 11.
- Viebrantz, A. F. P. M. (2013) “Bases de Patentes: Abordagem utilizando NoSQL e FULL-TEXT” Trabalho de Conclusão de Curso, Universidade Federal de Mato Grosso.

## Proposta de Redesenho do Catálogo *Online* da Biblioteca da UFMT

**Diego Vieira Neves, Igor de Oliveira, Adriano Roberto do Santos,  
Fernando Miranda Tanaka, Soraia Silva Prietch**

Curso de Sistemas de Informação – Universidade Federal de Mato Grosso (UFMT)  
Rodovia Rondonópolis-Guiratinga, MT 270 (KM 06) – Sagrada Família  
Rondonópolis - MT - Brasil

diegovineves@gmail.com, igor\_de\_oliveira@hotmail.com,  
adrianorsantos5@gmail.com, fmiranda50@globbo.com, soraia@ufmt.br

**Abstract.** *This paper presents a usability study regarding the online catalog provided by the Library of the Universidade Federal de Mato Grosso/Rondonopolis. In this study, the redesign of the system is proposed taking into account elements of the taxonomy of usability. The main result achieved was the high fidelity prototype development, which presents a list of modifications aiming at further improving the user interaction with the system.*

**Resumo.** *O presente trabalho apresenta um estudo de usabilidade realizado no catálogo online disponibilizado pela Biblioteca da Universidade Federal do Mato Grosso/Rondonópolis. Neste estudo, o redesenho do sistema é proposto tomando-se como base os elementos da taxonomia de usabilidade. Como resultado principal obteve-se o desenvolvimento de um protótipo de alta fidelidade que apresenta uma série de alterações visando a melhoria futura na interação dos usuários com o sistema.*

### 1. Introdução

O processo de automação das bibliotecas possibilitou aos usuários adquirir e obter um vasto número de informações de forma simples e eficaz, com o auxílio da Internet, as bibliotecas passaram a disponibilizar catálogos *online* que permitem a consulta dos itens bibliográficos oferecidos pela organização. A Biblioteca da Universidade Federal do Mato Grosso (UFMT) utiliza o software de gerenciamento de biblioteca Pergamum, que conta com um catálogo online, o qual pode ser acessado pela Internet, ampliando as condições de acesso pelos usuários. Conforme o Portal Pergamum (2014), esse é um sistema de gerenciamento de bibliotecas desenvolvido pela Divisão de Processamento de Dados da Pontifícia Universidade Católica do Paraná (PUC-PR) e atualmente é utilizado por aproximadamente 8000 bibliotecas no Brasil.

Os catálogos *online* possibilitam uma melhor interação com o usuário, entretanto conforme afirma Coelho (2006), é indispensável a preocupação com uma interface fácil de usar e intuitiva. Silveira (2003) expõe que os usuários não possuem as orientações de um bibliotecário nas consultas *online*. Deste modo segundo Silveira (2003), fazer com que o usuário compreenda melhor o uso do sistema e possa encontrar as respostas para as suas perguntas em um menor tempo possível são soluções que podem ser encontradas através da realização de estudos na área de Interação Humano-Computador (IHC).

Diante deste contexto, este trabalho realiza uma avaliação de usabilidade no referido catálogo *online*, disponibilizado pela Biblioteca da UFMT/Rondonópolis, com o objetivo de identificar as possíveis dificuldades dos usuários, bem como propor melhorias e a reformulação do atual catálogo disponibilizado pela instituição.

Este estudo está dividido em quatro seções. Na Seção 1 foi feita a introdução ao estudo realizado; na Seção 2 é apresentada a metodologia utilizada no desenvolvimento deste trabalho onde são descritas as etapas que envolveram esse processo; a Seção 3 expõe os resultados e considerações decorrentes da análise dos dados metodológicos obtidos durante o estudo de caso; por fim, na Seção 4 é realizado o fechamento do estudo em forma de conclusão.

## 2. Metodologia

O estudo foi desenvolvido durante a disciplina de Interface Homem-Máquina oferecida ao curso de Sistemas de informação da Universidade Federal do Mato Grosso/ Rondonópolis no ano letivo de 2013/2 e está dividido em duas etapas.

Na Etapa 1, inicialmente foi realizado um levantamento das possíveis deficiências apresentadas pelo catálogo *online*, seguindo os princípios e conceitos de taxonomia da usabilidade, proposto por Ferreira e Nunes (2008). Logo após foi feita a aplicação de um questionário aos usuários da Biblioteca da Universidade Federal do Mato Grosso/Rondonópolis. O questionário utilizado foi disponibilizado *online* durante 10 dias e apresentava 15 questões (14 fechadas e 01 aberta) com respostas não obrigatórias, pois caso o participante tivesse dificuldade ou não quisesse opinar o mesmo poderia deixar a questão em branco.

Com base nos dados obtidos, na Etapa 2 foi realizado um planejamento relativo às mudanças necessárias para alcançar melhorias do processo de prototipação. Logo após foi proposto o redesenho do catálogo *online* tendo como principal objetivo o desenvolvimento de uma interface mais interativa, possibilitando ao usuário uma melhor compreensão do conteúdo exposto pelo sistema. Nesta etapa, foram gerados: a lista dos requisitos não funcionais, várias versões de Wireframes<sup>1</sup>, e um protótipo navegável.

## 3. Resultados

### 3.1. Aplicação de questionários

Para realização da pesquisa, inicialmente, foram elaboradas 36 questões elucidadas por meio de *brainstorming*<sup>2</sup> entre os membros da equipe. No segundo momento foi realizada uma seleção entre as melhores questões proposta, resultando em um questionário piloto. Este questionário piloto foi aplicado a 04 usuários para identificar as principais dificuldades relativas à compreensão das questões. Após isso, foi realizada a reformulação do questionário piloto para o questionário efetivo, o qual foi disponibilizado *online* por meio do website Online Pesquisa<sup>3</sup> contendo 15 questões. Após a disponibilização do questionário efetivo<sup>4</sup>, foi realizada divulgação por meio de

---

1 Segundo Cybis, Betiol e Faust (2007), *Wireframes* são maquetes (em papel ou em meio digital), usados para sugerir a estrutura de um aplicativo, demonstrando os recursos que serão apresentados para o usuário.

2 Conforme descreve Cybis, Betiol e Faust (2007), *Brainstorming* é uma dinâmica utilizada para auxiliar um grupo de pessoas a criar o máximo de ideias no menor tempo possível.

3 Disponível em: <https://www.onlinepesquisa.com/>.

4 Disponível em: <http://goo.gl/5DAOVX>.

envio de e-mails e publicações em redes sociais. Por fim, a tabulação dos dados e a geração dos gráficos foram efetuadas para auxiliar a análise dos dados coletados pelo questionário efetivo.

A partir da descrição e análise dos dados, verificou-se a participação de 14 usuários, sendo estes 05 do sexo masculino e 09 do sexo feminino, os entrevistados apresentaram idades entre 23 a 51 anos de idade. Quanto ao grau de conhecimento em informática: 57,14% dos usuários disseram ter conhecimento básico; 28,57 afirmaram ter conhecimento médio; e, apenas 14,29% informaram que possuíam conhecimento avançado. Quando questionado sobre ter conhecimento da existência do recurso de catálogo *online* oferecido pela instituição: 92,86% responderam que sim; e, apenas 7,14% declaram que não. Foi identificado ainda que 50% dos usuários entrevistados já tiveram acesso ao treinamento oferecido pela instituição para o uso do catálogo *online*; 28,57% não sabiam da existência do treinamento; e, 21,43% informaram não ter recebido treinamento.

Referente ao conceito de acessibilidade (Questão 6), quando questionado se o acesso ao catálogo *online* pelo website da UFMT era de fácil identificação (localização): 57,14% afirmaram que sim; e, 42,86% dos usuários responderam que não. Verificou-se ainda que dentre a amostra entrevistada: 64,29% disseram utilizar o catálogo em outras localidades (diferentes contextos de uso) além da biblioteca, enquanto 35,71% responderam que não utilizam os serviços oferecidos em outras localidades. Dentre os entrevistados, 78,6% afirmaram que utilizam regularmente os serviços da biblioteca, e 21,4% disseram que raramente utilizam.

Quando questionado quais os tipos de filtros o usuário utilizava ao realizar uma busca no sistema (Questão 9), 50% afirmaram que realizam a busca usando o título do livro; 21,43% responderam que usam o nome do autor com parâmetros; 14,29 disseram que utilizam o filtro por assunto; e, 14,29 declaram não conhecer os recursos de filtro no sistema. Em relação ao acesso ao perfil do aluno (Questão 10), efetuando o *login* com usuário e senha: 85,71% responderam que já havia acessado; e, 14,29 afirmaram que nunca utilizaram o recurso. Foi verificado ainda que 64,29% dos entrevistados disseram que nunca tiveram dificuldades em efetuar o *login* no sistema e 35,71% declaram ter enfrentado dificuldades em realizar o *login*. Quando questionados sobre o serviço de recuperação de senha 84,62% disseram nunca ter utilizado o serviço e 15,38% afirmaram que precisaram do recurso, porém não obtiveram êxito.

Ao perguntar sobre a complexidade de interação e a compreensão do sistema: 64,29% dos usuários avaliaram como bom, 21,43% afirmaram ser regular, 7,14% classificaram como ruim e 7,14% consideram péssimo. No conceito de confiabilidade das informações retornada 85,71% dos entrevistados responderam que as informações são boas e 14,29% disseram ser regular.

No que se refere ao comportamento esperado pelo website: 61,54% dos participantes jugaram como regular; e, 38,48% atestaram ser bom. Em relação ao recurso de pesquisa oferecido pelo catálogo: 64,29% disseram ser bom o recurso apresentando pela ferramenta; 14,29% descreveram como regular; 14,29% avaliaram como ruim; e, 7,14% consideram como sendo ótimo. Ainda no conceito de pesquisa quando questionado sobre a eficiência dos filtros de busca: 42,86% declaram ser regular, 35,71% responderam como sendo bom; 14,29% classificaram como ótimo; e, 7,14% apontaram ser péssimo o recurso disponibilizado. No quesito de serviço de reserva e renovação oferecido pelo sistema, 35,71% declarou ser ótimo, 28,57% consideram bom, 28,57% responderam como sendo regular e 7,14 afirmou ser péssimo. Por fim, ao que se refere à consulta da situação do usuário na biblioteca: 35,71% avaliaram o recurso como ótimo; 35,71% descreveram como bom; 14,29% apontaram como sendo ruim; 7,14 avaliaram como regular; e, 7,14 classificaram como péssimo.

Como base nos dados obtidos, observou-se que o sistema mantém uma boa relação com seus usuários, entretanto, conforme pode ser evidenciado o sistema apresenta algumas falhas que necessitam ser corrigidas seguindo os princípios das Taxonomias de Usabilidades.

### 3.2. Levantamento dos requisitos conforme a taxonomia da usabilidade

Nesta etapa são apresentados os requisitos, baseada em itens da Taxonomia de Usabilidade<sup>5</sup>, propostos por Ferreira e Nunes (2008). Essa estrutura, segundo Pressman (2004) é composta de requisitos não funcionais que podem ser agrupadas em duas categorias: (1) *Exibição da informação*: consistência, *feedback*, nível de habilidade e comportamento do usuário, percepção humana e metáforas, minimização de carga de memória, eficiência no diálogo, classificação funcional dos comandos, manipulação direta, exibição exclusiva de informação relevante, uso de rótulos, uso adequado de janelas, projeto independente da resolução da tela; (2) *Entrada de dados*: mecanismo de ajuda, prevenção de erros, tratamento de erros.

Durante o estudo de usabilidade identificou-se, dentre outros problemas, que o sistema contém falhas no quesito de *consistência* (na categoria de exibição da informação); neste caso, o sistema faz o uso inadequado de cores e apresenta a falta de contraste adequado entre o fundo, campos e títulos, o que pode trazer dificuldade para usuários com algum tipo de deficiência visual. A análise descreve ainda inconsistência nos quesitos de *feedback*. No catálogo há ausência de uma melhor indicação para os filtros de pesquisas avançada utilizado. As buscas realizadas pelo sistema não trazem como prioridade o acervo do *campus* universitário em que o aluno que está matriculado (identificado ao efetuar o *login* no sistema). Além disso, verificou-se a falta de um mecanismo, que apresente ao usuário, exemplares similares à busca realizada retornando itens, por exemplo, do mesmo autor ou sobre o mesmo assunto.

O levantamento identificou também, uma precariedade nos requisitos de nível de habilidade e comportamento do usuário, o sistema não dispõe de recurso de recuperação e visualização de senha, o que pode gerar transtornos ao usuário. Outro requisito a ser mensurado refere-se ao recurso de áudio CAPTCHA utilizado na tela de *login* que durante o estudo apresentava idioma o inglês.

O estudo identificou também falha no quesito de percepção humana; o tamanho do texto e estilo de fontes utilizado pelo sistema não pode ser personalizado pelo usuário, dificultando a acessibilidade. Durante a análise foram identificadas ainda irregularidades no quesito de metáforas; o catálogo *online* apresenta inconsistência na utilização: (i) de ícones e funcionalidades do sistema como, por exemplo, o botão para filtros de pesquisas avançada apresentado pelo sistema; (ii) de um ícone que deveria ser associado a configurações e não apresenta nenhuma funcionalidade; e, (iii) do botão para ajustar fonte da página não realiza a funcionalidade devida abrindo uma janela de ajuda.

No que se refere à categoria de entrada de dados, a análise demonstrou que existe algumas falhas no mecanismo de ajuda, a opção existe, porém não retorna nenhuma funcionalidade na janela inicial. Tal funcionalidade é de extrema importância, pois possibilita que o usuário conheça a função de determinado item no sistema. Algumas falhas ainda foram identificadas quanto ao quesito de prevenção de erros; isso foi evidenciado na tela de cadastro e na tela de inserção de

---

<sup>5</sup> De acordo com Ferreira e Nunes (2008), a Taxonomia de Usabilidade foi proposta com o objetivo de sistematizar o processo de definição de requisitos (funcionais e não funcionais) e de minimizar as falhas de usabilidade em produtos ou sistemas computacionais.

informações pessoais, sendo que o sistema não faz nenhum tipo inibição de itens ou validação dos dados preenchidos, bem como não apresenta alertas para o usuário. Além disso, na opção de alterar o cadastro o botão gravar já vem habilitado e não apresenta nenhuma mensagem perguntando se o usuário deseja realmente alterar os dados.

### 3.3. Construção do protótipo

Nesta subseção são expostos os conceitos e processos envolvidos na prototipação do novo modelo de catálogo *online*, enfatizando a abordagem do desenvolvimento centralizando na necessidade do usuário, e os requisitos levantados a partir do estudo de usabilidade. Neste contexto, a Tabela 1 descreve os requisitos considerados durante o processo de prototipação.

**Tabela 1. Requisito a ser considerado durante a etapa de concepção do protótipo**

| FALHAS   | REQUISITOS NÃO FUNCIONAIS   |
|--|---|
| Alguns ícones não correspondem às funcionalidades do sistema | Reavaliar ícones utilizados e suas devidas funcionalidades  |
| Recursos de <i>login</i> / senha                             | Incluir recurso de recuperação e visualização de senha; Corrigir áudio do CAPTCHA que está em inglês.   |
| Busca  | Melhorar a indicação de filtro de pesquisas avançadas; Priorizar o <i>campus</i> universitário do usuário; Sugerir exemplares complementares ao assunto pesquisado. |
| Quesitos de acessibilidade                                   | Melhorar contraste das páginas.   |
| Ajuda  | Corrigir funcionalidade: o botão existe, porém não retorna nenhuma ação na janela inicial.  |
| Inibição dos itens   | Efetuar a inibição de itens não válidos; Apresentar alertas para o usuário.   |

Na sequência, com base nos dados coletados, foi possível desenvolver um protótipo de alta fidelidade. Neste sentido o catálogo *online* foi reformulado com base nos conceitos de taxonomia da usabilidade, abordado por Ferreira e Nunes (2008). Para o desenvolvimento do novo sistema foi utilizada a ferramenta Adobe Dreamweaver, sendo construído o total de 23 telas utilizando os recursos de HTML e CSS.

## 4. Conclusão

Como objetivo de trabalhos futuro pretende-se apresentar aos usuários o protótipo desenvolvido durante a realização deste trabalho, expondo as melhorias executada durante o processo de prototipagem e desta forma obter o *feedback* do novo modelo de catálogo *online*.

Conforme exposto anteriormente o sistema Pergamum é um software proprietário, neste contexto é importante ressaltar que o desenvolvimento do protótipo descrito neste estudo teve como principal objetivo compreender e consolidar os conceitos de IHC; visto que foi desenvolvido no âmbito de uma disciplina e não de um projeto em parceria com a empresa desenvolvedora. O objetivo do protótipo não é propor mudanças em um sistema proprietário, pois existem profissionais capacitados na empresa que desenvolveu o software. No entanto, é importante enfatizar que esse estudo gera contribuições que podem ser positivas à empresa. Sendo assim, o presente estudo se configura em um trabalho acadêmico, sem interesses e fins lucrativos. Entretanto, existe a ideia de encaminhar o resultado final à empresa responsável pelo desenvolvimento para que estas análises possam auxiliá-los a efetuar possíveis melhorias e correções no referido sistema.

Conforme pode ser evidenciado durante a análise deste estudo de caso, a utilização dos conceitos da usabilidade se mostra necessário para o desenvolvimento de um sistema. Deste modo a usabilidade de sites institucionais é um fator preponderante para que estes cumpram seus objetivos básicos de garantir o acesso a informação aos seus respectivos usuários.

No entanto, no decorrer do estudo não foi possível avaliar o protótipo junto aos usuários; esta etapa do ciclo de IHC fica como sugestão de trabalhos futuros, de modo a expor as melhorias executada durante o processo de identificação de falhas de usabilidade e de prototipação, bem como conhecer a opinião dos usuários a respeito das propostas de modificações.

Ainda, como sugestão trabalho futuro, pode-se realizar a reavaliação dos dados obtidos de forma interdisciplinar juntamente com o curso de Biblioteconomia da UFMT/Rondonópolis, assim como recomenda o trabalho de Dziekaniak (2004). Desta forma, com a colaboração de um bibliotecário (ou estudante de graduação), com afinidade do tema, potencializar e reforçar os resultados já apresentados. Neste aspecto, tornaria possível também identificar possíveis falhas não identificadas no decorrer deste trabalho, contribuindo na elaboração de um protótipo mais eficiente, especialmente, no que se refere à recuperação das informações por parte dos usuários.

## Referência

- Coelho, Luana Máira Plácido. (2006). Usabilidade de catálogos online: estudo exploratório dos OPACs das Universidades Públicas Paulistas (USP, UNESP, Unicamp). Monografia, Curso de Biblioteconomia e Documentação, Universidade de São Paulo, São Paulo, 2006. Disponível em: <<http://goo.gl/SPWMLq>>. Acesso em: 04/08/2014.
- Cybis, Walter; Betiol, Adriana H.; Faust, Richard. (2007). Ergonomia e Usabilidade: Conhecimentos, Métodos e Aplicações. São Paulo: Novatec Editora.
- Dziekaniak, Gisele V. (2004). Participação do bibliotecário na criação e planejamento de projetos de softwares: o envolvimento com a tecnologia da informação. Revista Digital de Biblioteconomia e Ciência da Informação (RDBCI), v. 2, n. 1, p. 37-56.
- Ferreira, Simone B. L.; Nunes, Ricardo R. (2008). e-Usabilidade. Rio de Janeiro: LTC.
- Portal Pergamum. (2014). Pergamum Sistema Integrado de Bibliotecas: Conheça o Pergamum. Disponível em: <<http://rbbd.febab.org.br/rbbd/article/view/104>>. Acesso em: 22/08/14.
- Pressman, Roger S. (2004). Engenharia de Software. 6ª ed. São Paulo: McGraw-Hill.
- Silveira, Alex da. (2003). A usabilidade em catálogos online. Monografia, Curso de Biblioteconomia, Escola de Biblioteconomia da Universidade do Rio do Janeiro, Rio de Janeiro, 2003 Disponível em: <<http://goo.gl/mr5kfz>>. Acesso em: 08/08/2014.

# COMPARAÇÃO DE METAHEURÍSTICAS PARA O PROBLEMA DE CLUSTERIZAÇÃO AUTOMÁTICA

Alexandre Tomasi<sup>1</sup>, Luciana Lee<sup>2</sup>

<sup>1</sup>Universidade Federal de Mato Grosso (UFMT) - Instituto de Ciências Exatas e da Terra

<sup>2</sup>Departamento de Computação e Eletrônica - Centro Universitário Norte do Espírito Santo - Universidade Federal do Espírito Santo.

alexandretomasi18@gmail.com, luciana.lee@ufes.br

**Abstract.** In this paper we study different versions of a hybrid algorithm based on GRASP to solve the Automatic Clustering Problem. We studied a algorithm presented by [Cruz 2010] and we proposed some improvements in the solution. To test the algorithms, we used a set of instances generated using points on the  $R^2$  space. To analyze the performance of the algorithms we compared the values of the objective function of the obtained solutions and the execution time of the algorithms.

**Resumo.** Neste trabalho, são estudadas diferentes versões de um algoritmo híbrido baseado em GRASP para resolver o Problema de Clusterização Automática. É estudado um algoritmo apresentado por [Cruz 2010] e é proposto algumas melhorias na solução. Para testar os algoritmos, foram gerados casos de testes utilizando pontos no espaço  $R^2$ . Para analisar o desempenho dos algoritmos foram comparados os valores da função objetivo das soluções obtidas e o tempo de execução dos algoritmos.

## 1. Introdução

A clusterização (ou agrupamento) é um processo no qual objetos com alto grau de similaridade são reunidos em um mesmo grupo. Este é denominado de *cluster*. Estes agrupamentos fornecem estratégias para uma melhor solução na organização ou análise de um conjunto de objetos (dados). Quando a quantidade de *clusters* é previamente definida o problema é conhecido como *Problema de Clusterização* (PC). Quando não se conhece o número de *clusters* para a solução ótima, o problema é conhecido como *Problema de Clusterização Automática* (PCA). Neste trabalho, o segundo problema será abordado.

O PCA, por não se saber o número ideal de *clusters*, é um problema de otimização de elevada complexidade. O objetivo deste trabalho é estudar metaheurísticas para encontrar soluções para o PCA. Para isso, foram estudados alguns algoritmos apresentados em [Cruz 2010] para tratar o PCA. Um dos algoritmos estudados foi uma versão híbrida do algoritmo GRASP com o algoritmo VNS. Ao analisar tal algoritmo, observou-se que algumas melhorias podem ser implementadas de forma a obter soluções de melhor qualidade e em menor tempo de execução. Sendo assim, o algoritmo foi implementado e melhorado em dois pontos, que serão

apresentados na seção 3. Na Seção 4, o algoritmo proposto por [Cruz 2010] e as versões com as melhoras foram comparadas quanto à qualidade das soluções. E na seção 5, têm-se as conclusões e trabalhos futuros.

## 2. Problema de Clusterização Automática

Existe uma grande variedade de aplicações e diferentes problemas para o PC, o qual é mais discutido na literatura do que para o PCA. Os métodos utilizados para o PCA, geralmente, são aproximados ou heurísticas. Não existe um algoritmo ideal para todos os problemas e, se tratando de metaheurísticas para PCA, se destacam os seguintes trabalhos: [Soares 2004], [Cruz 2010], [Semaan *et al* 2012].

A clusterização é aplicada em vários problemas reais, seja para processar imagens, customizar estratégias de marketing, aperfeiçoar configurações de máquinas em linhas de montagem, escalonar tarefas em ambientes multiprocessados, no emparelhamento de genes, ou como auxílio no diagnóstico de doenças [Soares 2004].

## 3. Algoritmos Implementados

Foi estudado o algoritmo GRASP híbrido proposto por [Cruz 2010]. Trata-se de uma metaheurística híbrida composta de duas fases: a construção da solução inicial e a busca local. A primeira fase trata-se da construção inicial do algoritmo GRASP e a segunda fase, a busca local, consistindo da metaheurística VNS utilizando as vizinhanças: *Inversão individual*, *Reconexão por Caminhos*, *Troca entre Pares*.

Além de implementar o algoritmo, foram feitas duas modificações na tentativa de melhorar as soluções. Uma modificação na fase de construção, chamado de *Gerar Clusters Parciais Densos (GCPD)* e um novo procedimento ao final do algoritmo, chamado de *Mudança de Clusters para Pontos Distantes (MCPD)*.

### 3.1 Fase de Construção dos Algoritmos

A etapa de construção tem como objetivo gerar uma solução inicial de boa qualidade para depois ser executada a busca local. Na etapa de construção, dois processos são comuns aos algoritmos estudados, propostos por [Cruz 2010], são eles: *procedimento Gerar Clusters Parciais (GCP)* e *Junção de Clusters Parciais Adjacentes (JCPA)*. O procedimento GCP irá criar os *clusters* parciais, e o procedimento JCPA fará com que diminua a quantidade desses *clusters* juntando os *clusters* parciais muito próximos.

Os *clusters parciais* são criados da seguinte forma: é selecionado aleatoriamente um ponto e todos os pontos ao redor deste, num raio de alcance  $r$ , que formarão junto a ele um *cluster* Figura 1. Isso é repetido até que todos os pontos estejam mapeados a algum *cluster*.

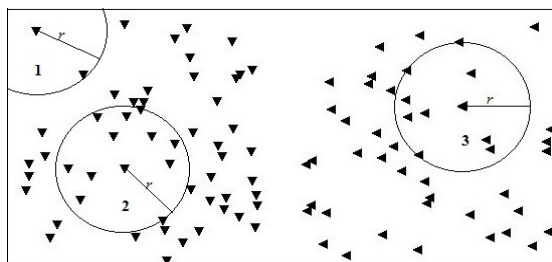


Figura 1. Processo de geração dos clusters parciais.

Após a geração dos *clusters* parciais os mesmos são melhorados, agrupando os *clusters* que estão muito próximos um do outro e, assim, diminuindo a quantidade de *clusters*, esse processo é realizado no procedimento JCPA.

No procedimento JCPA, os *clusters* que possuem 4 pontos ou menos, são considerados *clusters parciais pequenos*, já os *clusters* que possuem 10 ou mais pontos, são considerados *clusters parciais grandes*. Existem duas condições para que um *cluster parcial pequeno* se junte a outro *cluster*. A primeira condição é que um *cluster parcial pequeno* se juntará apenas a um *cluster parcial grande*. A segunda condição é que deve existir um ponto no *cluster parcial pequeno* com distância inferior a uma preestabelecida a um ponto do *cluster grande*. Depois de juntar os *clusters* próximos, o procedimento retorna os *clusters* iniciais.

### 3.1.1 Novo Procedimento

Ao analisar o processo GCP, foi observado que o mesmo gerava muitos *clusters* parciais, de tal forma que havia *clusters* parciais com poucos pontos. Sendo que, muitas vezes, a solução inicial poderia ficar melhor se os *clusters* parciais com poucos pontos fossem agrupados, formando um *cluster* apenas. Como o procedimento JCPA possui alguns requisitos para junção de *clusters*, muitos destes *clusters* não são unidos. Os mesmos ficavam por conta da Busca Local.

Para diminuir a quantidade de *clusters* parciais com poucos pontos, e gerar *clusters* parciais com mais pontos, foi feita uma modificação no procedimento GCP. Em vez de escolher um ponto aleatoriamente, e selecionar os pontos que estão a uma distância menor que  $r$ , para formar um *cluster* parcial, é escolhido um ponto verificando se o mesmo possui bastantes pontos ao seu redor, com uma distância menor que  $r$ , para se tornar um *cluster* parcial.

Sendo assim, os pontos que possuem mais pontos ao redor são selecionados para formar um *cluster* parcial, juntamente com os pontos próximos que possuem uma distância menor que  $r$ . Formando assim *clusters* parciais com maior número de pontos. O procedimento GCP modificado é chamado de *Gerar Clusters Parciais Densos (GCPD)*.

### 3.1.2 Representação da Solução

Após a etapa de construção, os *clusters* gerados  $B = \{ B_1, B_2, \dots, B_p \}$  serão os *clusters* iniciais. A solução é representada em um vetor binário com  $p$  posições. Cada posição desse vetor representa um *cluster* gerado na etapa de construção. Um *cluster* em  $B$  pode ser classificado como *cluster* pai ou *cluster* filho. O primeiro, será mantido na solução final, podendo ter outros *clusters* agrupados a ele. Um *cluster* filho será agrupado a um *cluster* pai e deixará de existir como *cluster* na solução final. No processamento da solução final, um *cluster* filho será agrupado ao *cluster* pai mais próximo. Na representação binária da solução, um *cluster* pai é representado pelo valor 1 e um *cluster* filho pelo valor 0. Como critério para saber qual *cluster* pai está mais próximo de um *cluster* filho, para sua junção, é utilizado a distância entre os centroides.

Por exemplo, se uma solução for com  $p = 5$  terá um vetor com 5 posições com zeros ou uns. Exemplo  $\{01001\}$ . Sendo assim, um vetor que possui 5 posições com 2

delas com valores 1, então a solução possuirá 2 *clusters* pais e 3 *clusters* filhos. Após a junção dos *clusters* filhos com os *clusters* pais, esse vetor representará uma solução com apenas 2 *clusters*. Então os *clusters* finais são criados. Suponha que  $C = \{C_1, C_2, \dots, C_k\}$  seja o conjunto de *clusters* finais, então  $C$  é uma solução para o problema.

### 3.1.3 Lista Restrita de Candidatos

Após a criação dos *clusters* parciais, ainda na fase de construção, é criada a Lista Restrita de Candidatos (LRC).

Para preencher a LRC, é gerada várias soluções (representação binária) que podem ter o mesmo número de pais ou número de pais diferentes. Por exemplo: 5 pais {111101}, 4 pais {101101}, 4 pais {011110} e assim por diante. Os *clusters* que serão pais são selecionados de forma aleatória.

Cada solução gerada é avaliada usando uma função objetivo (FO), de modo a serem ordenadas de forma não crescente de valor de função objetivo, pois para cada número de pais há um intervalo possível para o valor da FO, ou seja, um valor máximo e um valor mínimo para a FO. Para cada solução são armazenados o número de pais e o valor resultante da FO na LRC.

Como a LRC armazena quantias de pais que levam a soluções boas, esse é um fator guloso do algoritmo, sendo assim, após a criação da LRC, já é possível ter uma estimativa do número ideal de *clusters*.

### 3.2 Função de Avaliação

Existe um modelo para avaliar a qualidade de uma solução que não necessita saber, inicialmente, a quantidade de *clusters*. Este modelo é chamado de Índice Silhueta e foi proposto por [Kaufman and Rousseeuw 1990].

O modelo Índice Silhueta consiste em avaliar todos os elementos de uma solução. Isto é, se os elementos de um *cluster* são similares entre si e se os elementos de um *cluster* são diferentes dos elementos de outros *clusters*. Quanto mais os elementos forem similares no seu próprio *cluster* e quanto mais diferentes esses elementos são dos outros *clusters*, maior será o valor resultante desse modelo.

### 3.3 Fase de Busca Local

Na fase de busca local, é escolhido aleatoriamente um elemento da LRC. E com a quantidade de *clusters* pais desse elemento, é gerada uma solução inicial. Essa solução inicial é gerada de forma aleatória, ou seja, escolhendo aleatoriamente no vetor que representa a solução, as posições que representarão os *clusters* pais. Após gerar a solução inicial, essa passará pela busca local. O algoritmo GRASP Híbrido proposto por [Cruz 2010], utiliza três processos de busca local. São eles: o primeiro a *Inversão Individual*, o segundo a *Reconexão de Caminhos* e o terceiro a *Troca Entre Pares*.

#### 3.3.1 O Procedimento Mudança de *Clusters* para Pontos Distantes (MCPD)

Foi observado que os algoritmos, ao serem executados em uma massa de dados com pontos muito espalhados, não conseguiam gerar bons resultados. O algoritmo gerava *clusters* com pontos distantes do seu centro, tão distantes que poderiam ser melhores agrupados em outro *cluster*, de modo a melhorar a solução final.

Como a etapa de construção gera apenas *clusters* iniciais e a etapa de busca local faz uma junção dos *clusters*, em nenhum momento é analisado se pode fazer mudanças em apenas um ponto, que possa trazer benefícios à solução final. Tendo em vista tal fato, foi proposto um processo que é executado após o término da busca local do algoritmo, que faz uma busca desses pontos na solução final, para que os mesmos possam fazer parte de outro *cluster* mais próximo, melhorando assim a solução.

#### 4 Resultados

Neste trabalho foram utilizados pontos no espaço  $R^2$  como entradas para o programa. A quantidade de pontos de cada teste varia de 100 a 500, podendo assim avaliar os algoritmos em diferentes massas de testes.

Utilizou-se no trabalho como massa de testes, pontos no primeiro quadrante do plano cartesiano e com valores das coordenadas  $x$  e  $y$  variando de  $[0,700]$ . Para cada teste foram criadas 10 instâncias diferentes para testar os algoritmos. As instâncias de pontos foram gravadas em arquivos para que na comparação dos algoritmos e na etapa de calibração, os mesmos tenham como entrada as mesmas instâncias. Não existem pontos repetidos. Os algoritmos foram executados duas vezes, sendo feito a média de tempo e silhueta desses resultados.

Foram construídas cinco tipos de instâncias: a primeira foi escolhendo pontos aleatoriamente; a segunda foi com pontos distribuídos a formar 4 clusters; a terceira foi com pontos distribuídos a formar 9 clusters; a quarta foi com pontos distribuídos a formar 16 clusters; a quinta foi com pontos distribuídos a formar 25 clusters. O gráfico com os resultados é apresentado na Figura 2.

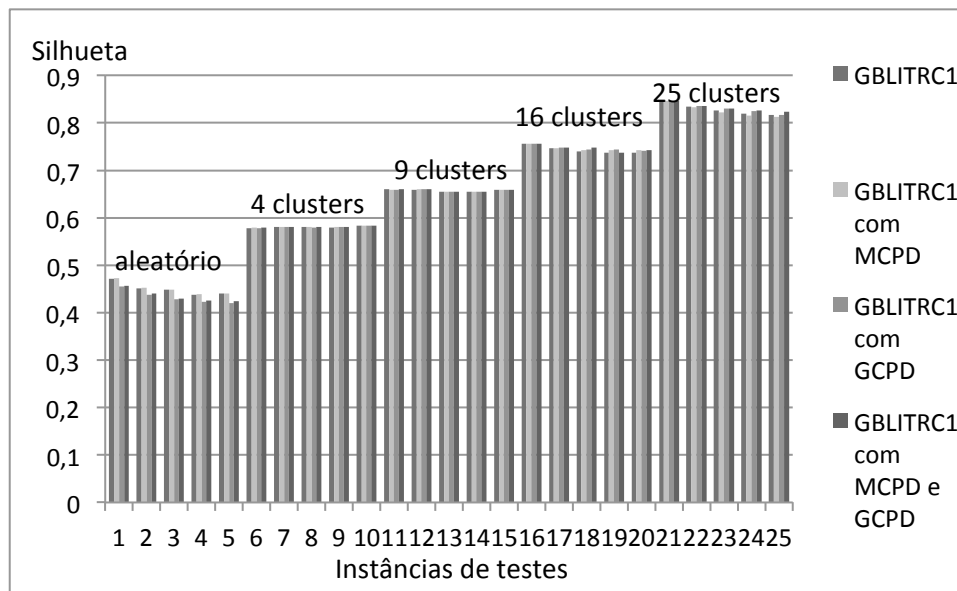


Figura 2. Comparação das silhuetas entre os algoritmos.

Os algoritmos foram comparados, analisando os resultados da inserção dos procedimentos GCPD e MCPD. O procedimento *GCPD* gerou grandes melhoras nos tempos de execução dos algoritmos, aproximadamente 20%, possivelmente por gerar menos *clusters* parciais. Mas para instâncias de testes com pontos aleatórios, o valor resultante das silhuetas obteve perdas significativas, aproximadamente 4%. Como instâncias com pontos aleatórios tem poucas ou nenhuma aglomeração de pontos, o

procedimento não tem um bom desempenho. Já nas instâncias que possuem aglomerações, os resultados foram muito semelhantes, com até melhoras em certos resultados. Nos testes usando o procedimento *MCPD*, o algoritmo obteve pequenos ganhos em quase metade dos resultados, sem acarretar em perdas significativas de tempo.

## 5 Conclusão

Neste trabalho foram estudadas metaheurísticas para o PCA. Sendo mais a fundo o trabalho de [Cruz 2010], o qual trabalha com algumas metaheurísticas. Ao analisar o algoritmo GRASP, foi observado que sua versão demanda muito tempo de execução. Sendo assim, foi proposto uma modificação na etapa de construção, gerando um novo procedimento chamado de *GCPD*. Com objetivo de gerar um bom número de *clusters* parciais. Melhorando assim o tempo de execução das buscas locais acarretando em grandes ganhos de tempo para os algoritmos.

Ao executar os algoritmos em massas de testes, principalmente nas que possuíam pontos muito espalhados, observou-se que muitas vezes eram gerados *clusters* com algum ponto muito distante do seu centroide. Portanto, foi proposto um novo procedimento, a ser executado no fim do algoritmo, chamado de *MCPD*. Com o objetivo de analisar esses pontos, verificando se podem ser melhor agrupados em outros *clusters* próximos, na tentativa de obter melhores resultados, sem grande custo de tempo de execução.

É de interesse como trabalhos futuros, desenvolver outras versões do algoritmo GRASP utilizando diferentes extensões da metaheurística VNS. E também desenvolver outras metaheurísticas como *Busca Tabu* para verificar seu desempenho ou até mesmo desenvolver versões híbridas com GRASP.

## 6. Referências Bibliográficas

- Cruz, M. D. (2010) “O Problema de Clusterização Automática”. p. 120.
- Glover, F. (1996) “Tabu Search and Adaptive Memory Programming — Advances, Applications and Challenges”. 7, p. 75.
- Kaufman, L., and Rousseeuw, P. J. (1990) “*Finding Groups in Data An Introduction to Cluster Analysis*”, New Jersey, United States of America: WILEY-INTERSCIENCE.
- Semaan, G. S., Cruz, M. D., Brito, J. a., and Ochi, L. S. (2012) “Proposta de um Método de Classificação Baseado em Densidade para a Determinação do Número Ideal de Grupos em Problemas de Clusterização”, *Journal of the Brazilian Computational Intelligence Society*, 10, pp. 242-262.
- Soares, S. S. (2004) “Metaheurísticas para o Problema de Clusterização Automática”. p. 137.

# MODEFS: Towards a Lightweight Model-Driven Environment for Formal Software Engineering

Elias Adriano Nogueira da Silva<sup>1,3</sup>, Valdemar Vicente Graciano Neto<sup>1,2</sup>

<sup>1</sup>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

<sup>2</sup>Instituto de Informática – Universidade Federal de Goiás (UFG)  
Alameda Palmeiras, Quadra D, Câmpus Samambaia  
P.O Box 131 - CEP 74001-970 - Goiânia - GO - Brazil

<sup>3</sup>Universidade Nova de Lisboa - Faculdade de Ciências e Tecnologia  
Quinta da Torre, Campus Universitário, CEP 2829-516. Caparica, Portugal

eliasnog@icmc.usp.br, valdemarneto@inf.ufg.br

**Abstract.** *Formal methods (FM) allow unambiguous software specifications, and Model-Driven Engineering (MDE) lives its golden-age as a software engineering paradigm that transforms abstract models into source code and other software artifacts. MDE can raise the level of abstraction for FM-based approaches by means of models. Such association creates a lightweight model-driven formal method and improves communication between developers, providing communication alternatives for customer and other stakeholders, and automating verification of formal models, minimizing time and costs. Although the envisioned benefits are clear, tools, approaches, and methods to integrate MDE and FM have not emerged. Thus, as a first effort towards a complete environment for integrating both paradigms, this paper presents an architectural proposal for MODEFS project, a lightweight MOdel-Driven Environment for Formal Specification.*

## 1. Introduction

Formal Methods (FM) are being increasingly widespread in Software Engineering because they enhance safety, reliability and quality of critical software systems. Furthermore, they have been used in non-critical software as a way to ensure software quality. At this direction, Model-Driven Engineering (MDE) is a software development paradigm which can be associated to FM by means of abstract models that can increase level of abstraction in FM approach, automatically transforming abstract formal models in code and other artifacts. Additionally, traditional FM tools for formal specification, and formal validation can be associated to this process to emerge a lightweight model-driven formal method. However, an integrated and effective tool support is necessary to support such association, and promoting a wide adoption of FM in Software Engineering processes.

This paper presents MODEFS, a lightweight MOdel-Driven Environment for Formal Specification which composes and integrates FM tools, methods, to repositories, and

Domain-Specific Languages to create an environment for automatic generation of software based on Model-Driven Engineering (MDE) and FM. The main contribution of this paper is a public architectural proposal of such environment, highlighting its major parts and how they interoperate. We choose a specific viewpoint to show for space limitations. Additionally, we built a prototype and a proof of concept implemented as an Eclipse <sup>1</sup> plug-in.

## 2. Model-Driven Engineering and Formal Methods

Software development techniques have been improved. However, concerns about systems development, modeling, reuse, productivity, maintenance, documentation, validation, optimization, portability and interoperability tools are still recurrent.

Model-Driven Engineering (MDE) emerged aiming to solve some of these problems [Kleppe et al. 2003] shifting focus of modern development methodologies from implementation technology to a conceptual modeling perspective [Costa et al. 2010]. In this context, models are first-class citizen [Sendall and Kozaczynski 2003, Amrani et al. 2012], and transformation mechanisms are used to generate code from abstract models, which reduces the efforts of developers [Kleppe et al. 2003] and increase portability and productivity [da Silva et al. 2013].

On the other hand, FM are recognized as a solution to problems in software engineering as quality assurance, traceability, correctness, consistency, and others, and there is an increased recognition that a broad adoption of FM requires they must be lightweight [Agerholm and Larsen 1999, Breen 2005]. Additionally, software engineering environments should support FM without needing to learn arcane mathematical notations or subtle theoretical concepts, and a successful method must also support requirements specification as an activity [Breen 2005]. Formal methods have been used to construct models without ambiguity and inconsistencies [Valles-Barajas 2012].

Model finders are an essential class of software that are used in Formal Methods approach. Given a set of logical and mathematical statements, a Model Finder is capable to verify if the set of statements is consistent, i.e, if statements are not contradictory between each other. A Model Finder can ensure that the statements are logically consistent (they are satisfiable or SAT) or if there is no a way where that set of statements could be consistent(the are unsatisfiable or UNSAT). One of the most famous Model Finders is Z3 [De Moura and Bjørner 2008], developed by Microsoft Research in USA (An online version is available <sup>2</sup>). A Model Finder can work under different theories. Z3 is based on satisfiability modulo theory (SMT). Solvers for SMTs are concerned with the problem of determining the satisfiability of a set of formulas in some first order theory [Reynolds et al. 2013]. A model can be obtained by integrating a finite model finding capability into the architecture of a modern SMT solver [Reynolds et al. 2013].

We have perceived *Formal Methods* (FM) and *Model-Driven* do not often appear in the same sentence [Graciano Neto 2014]. MDE could be applied to FM and vice versa [Gargantini et al. 2010], but there is a lack of tools that implement these ideas, what can be critical once the adoption of a technology uses to depend intensively of avail-

---

<sup>1</sup><http://www.eclipse.org/>

<sup>2</sup><http://z3.codeplex.com/>

able tools to use [Hutchinson et al. 2011]. Representing software as models is an excellent idea, but there is no guarantees about its correctness, conformance, quality or rigor [Graciano Neto 2014]. On the other hand, applying formal specification and validation is worth, but it is not so abstract as the models used in software engineering industry. FM allow a software engineer to create a more complete, consistent, and unambiguous specification. Then, an association between these concepts and an associated tool support can bridge a gap on both sides, exploring their inherent complementary weaknesses.

### 3. A lightweight MODEL-Driven Environment for Formal Specification (MODEFS)

MODEFS is an environment to support formal specification and validation of software using models in high level of abstraction, bridging the gap between formal specifications and code through model transformations.

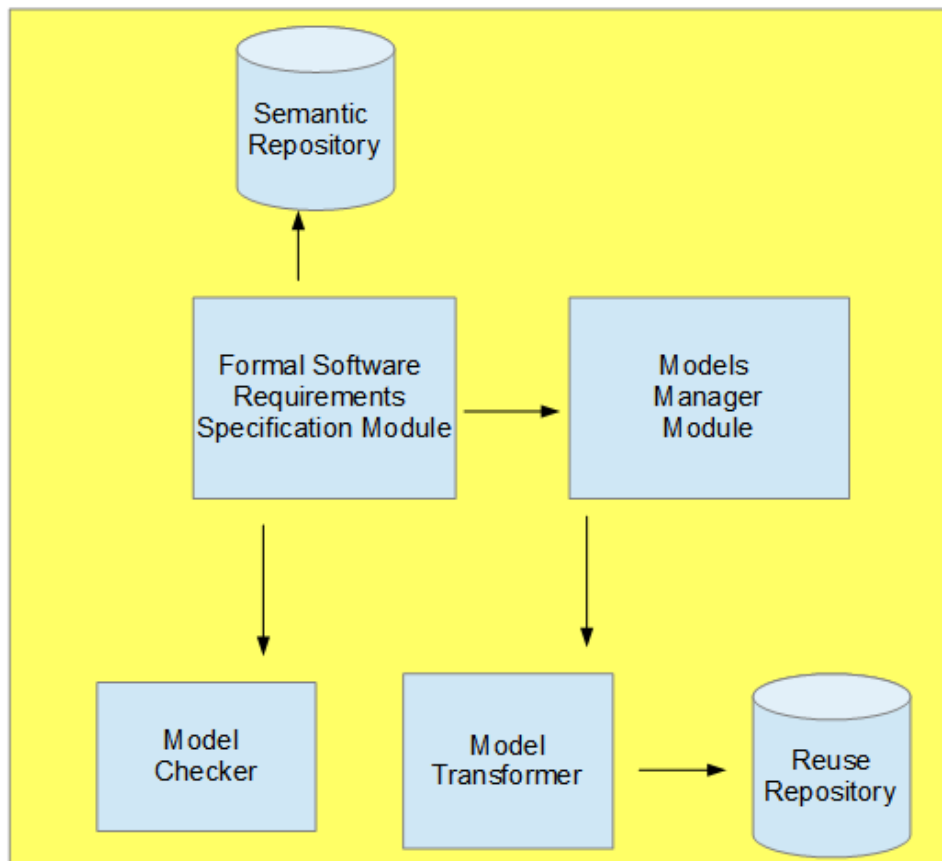


Figure 1. MODEFS: an structural view.

MODEFS infrastructure architecture is composed by these modules: **Semantic Repository**, **Reuse Repository**, **Formal Software Requirements Specification Module**, **Models Manager Module**, **Model Prover**, and **Model Transformer**, as depicted in Figure 1. We chose to depict an MODEFS architecture under only one view, equivalent to one of those proposed by Kruchten [Kruchten 1995].

Two repositories are used. Semantic Repository stores language structures to support the automatic identification of requirements could be converted to logic propositions

that could be validated through a Model Prover, as we implemented inspired in Z3. A structured and delimited portuguese must be used to specify requirements. An electronic dictionary as WordNet [Gomes et al. 2013] automatically classifies words according to its potential grammar class and subsequently identifying potential logic prepositions structured in a form *Subject Verb Predicate*.

Formal Software Requirements Specification Module uses services provided by the repositories. This module is responsible for supporting the software requirements specification. An environment is provided to specify software requirements in a structured way. After specification, services are provided by the Semantic Repository to classify identified words. Those services make it possible to identify phrases that are offered to users in such a way they could confirm if it is really a potential logic preposition that will compose the formal specification.

Model Prover module embodies two integrated tools: a Model Finder (as Z3) and a Model Checker (as NuSMV). Model Finder solves the satisfiability problem to assure the set of propositions have logic values assigned to them in a way that the conjunction between them are true. Model Checker is responsible to check consistency between those logic prepositions elicited. Consistency, ambiguity, and redundancy are automatically verified using a Model Checker.

Components will be used according to a restrict domain. For example, if a software is being developed to a home security system, words as *alarm*, and *invasion* could be associated to COTS components that could be composed in a transformation process to generate a software product.

Models Manager Module manages models generated by Formal Software Requirements Specification Module, as the formal software requirements specification and other models derived from that one in a more concrete level. And finally, Model Transformer module holds a model transformer component which is used to automatically generate software from formal models

#### **4. Proof of Concept for an Architecture Module**

As a part of MODEFS environment, a Model Finder was implemented to integrate the Model Prover component that we called ZEclipse. The Model Finder engineering was divided in two phases: analyser engineering (the component responsible for lexical, syntactic and semantic analysis), and the SMT solver component engineering.

##### **4.1. ZEclipse - A Model Finder Component**

Model-Driven Engineering (MDE) techniques and tool were used to support the ZEclipse engineering. Xtext<sup>3</sup> was used as an MDE tool to automatically generate an analyser for our model finder. Xtext is a model-driven framework for development of programming languages and Domain Specific Languages. In the context, we also implemented a semantic checker.

##### **4.2. SMT Solver Component**

SMT Solver follows the classic approach for satisfiability checking: Tableau Method [Keller and Heymans 2008]. A Tableau can be represented as a tree data structure. The

---

<sup>3</sup><http://www.eclipse.org/Xtext/>

root stores a conjunction of axioms that will be tested about their satisfiability. There are specific rules of expanding the tree. In our Tableaux, the axioms are expressed in First Order Logic and it is restricted to logic formulas. We do not solve mathematical problems as Z3 does. A set of deduction rules is used to construct the tableau.

During the solution of Tableau, priority is given to the solution of axioms with "implies" and "and" operations. "Or" operations are solved as late as possible since they generate the opening of Tableau branches and branches can open indefinitely. If the opening is unavoidable, it is done.

## 5. Conclusions and Future Work

This paper presented a public architecture of an environment to support a lightweight formal specification and validation for software based on Model-Driven Engineering principles, which is the main contribution. An Eclipse-based plugin was implemented as a prototype to perform coding-time verifications of a Z3 model specification as a proof of concept for an specific module provided in environment architecture proposal. Requirements Engineering support can be improved by increasing the controlled vocabulary to encompass linguistic equivalences to support temporal logic and other types of logic. Additionally, other common complementary models such as Software Requirements Specifications, UML class diagrams, use cases, and activity diagrams could be added to the SRS software specification. Additionally, this is a first step towards effectively bringing formal methods (FM) to software engineering tools reality. We believe that our contributions can foster the adoption of FM once make it easy to use FM in a software specification and the automatic software generation from abstract models.

## 6. Acknowledgments

The authors would like to thank FAPESP process number 2012/244887-3 for partially funding this research.

## References

- Agerholm, S. and Larsen, P. (1999). A lightweight approach to formal methods. In Hutter, D., Stephan, W., Traverso, P., and Ullmann, M., editors, *Applied Formal Methods — FM-Trends 98*, volume 1641 of *Lecture Notes in Computer Science*, pages 168–183. Springer Berlin Heidelberg.
- 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.
- Breen, M. (2005). Experience of using a lightweight formal specification method for a commercial embedded system product line. *Requirements Engineering*, 10(2):161–172.
- Costa, S. L., Graciano Neto, V. V., Loja, L. F. B., and de Oliveira, J. L. (2010). A Meta-model for Automatic Generation of Enterprise Information Systems. In *Proceedings of the First Brazilian Workshop on Model Driven Development (BW-MDD)*, BW-MDD '10, pages 45–52, Salvador, BA, Brazil. UFBA.

- da Silva, E. A. N., Fortes, R. P. M., and Lucredio, D. (2013). A Model-Driven Approach for Promoting Cloud PaaS Portability. In *Anual International Conference on Software Engineering-CASCON*.
- De Moura, L. and Bjørner, N. (2008). Z3: An Efficient SMT Solver. In *Proceedings of the Theory and Practice of Software, 14th International Conference on Tools and Algorithms for the Construction and Analysis of Systems*, pages 337–340, Berlin, Heidelberg. Springer-Verlag.
- Gargantini, A., Riccobene, E., and Scandurra, P. (2010). Combining formal methods and mde techniques for model-driven system design and analysis. *International Journal On Advances in Software*, 1:1–18.
- Gomes, M. M., Beltrame, W., and Cury, D. (2013). Automatic Construction of Brazilian Portuguese WordNet. In *X Encontro Nacional de Inteligência Artificial e Computacional (ENIAC)*, pages 1–10. UFCE.
- Graciano Neto, V. V. (2014). Model-Driven Development and Formal Methods: A Literature Review. In *Proceedings of the XI Annual Meeting of Computing, ENACOMP '14*, pages 1–10.
- Hutchinson, J., Whittle, J., Rouncefield, M., and Kristoffersen, S. (2011). Empirical assessment of mde in industry. In *Proceedings of the 33rd International Conference on Software Engineering, ICSE '11*, pages 471–480, New York, NY, USA. ACM.
- Keller, U. and Heymans, S. (2008). The sat-tableau calculus. In *Proceedings of the 21st International Workshop on Description Logics (DL2008), Dresden, Germany*. CEUR-WS.org.
- Kleppe, A., Jos, W., and Wim, B. (2003). MDA Explained, The Model-Driven Architecture: Practice and Promise. Addison-Wesley.
- Kruchten, P. (1995). The 4+1 view model of architecture. *IEEE Softw.*, 12(6):42–50.
- Reynolds, A., Tinelli, C., Goel, A., Krstic, S., Deters, M., and Barrett, C. (2013). Quantifier Instantiation Techniques for Finite Model Finding in SMT. In *CADE-24 – the 24th International Conference on Automated Deduction, Lecture Notes in Computer Science*, pages 377–391. Springer.
- Sendall, S. and Kozaczynski, W. (2003). Model transformation: The heart and soul of model-driven software development. *IEEE Software*, 20(5):42–45.
- Valles-Barajas, F. (2012). Using Lightweight Formal Methods to Model Class and Object Diagrams. *Computer Science and Information Systems*, 9(1):411–429.

# Arquitetura Inception: Uma solução de abertura de portas dinâmicas para o protocolo SMIP

Eduardo Nascimento de Souza Rolim<sup>1</sup>, Claudio de Castro Monteiro<sup>1</sup>, Fabiana Ferreira Cardoso<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia do Tocantins (IFTO)  
Palmas – TO – Brasil  
Grupo de Pesquisa em Redes de Computadores - GREDES

vndmtrx@gmail.com, ccm@ifto.edu.br, fabiana@ifto.edu.br

**Resumo.** *Esta proposta é focada em facilitar a implantação do protocolo SMIP para mobilidade de redes. Através do desenvolvimento de uma infra-estrutura de atravessamento de redes privadas envolvendo o modelo Cliente-Servidor que, instalado atrás de um dispositivo tradutor de endereços de rede (NAT), usa protocolos de abertura e mapeamento de portas dinâmicas para efetuar a sincronização das portas e permitir a comunicação entre o cliente e o servidor. Palavras-chave: SMIP, Atravessamento de NAT, STUN, Mobilidade de Redes.*

**Abstract.** *This proposal is directed to ease the deployment of the SMIP protocol for network mobility. Through the development of a infrastructure for private network transversal involving the client-server model, deployed behind Network Address Translation gateways, uses port opening protocols and dynamic port assignment to synchronize the port discovery and subsequent connection between the client and server.*

*Keywords: SMIP, NAT Transversal, STUN, Network Mobility.*

## 1. Introdução

Apesar de pouco uso em sistemas comerciais e em pesquisas relacionadas com a interconexão de redes, o atravessamento dinâmico de redes, definido através do protocolo desenvolvido em [Rosenberg et al. 2003], é um recurso amplamente utilizado em redes ponto-a-ponto, sob a forma de protocolos de descoberta de rede, permitindo a interligação entre clientes que se encontram atrás de redes mascaradas.

Este trabalho se propõe a desenvolver a Arquitetura *Inception*, um complemento à arquitetura do protocolo *SMIP* através do uso de protocolos de abertura e gerenciamento de portas dinâmicas em dispositivos tradutores de nomes de rede (*NAT*), aumentando a capacidade de descoberta dos agentes do *SMIP*. Tal proposta se faz através da eliminação da necessidade de se ter um *IP* público nos agentes e da adição de uma entidade responsável por informar da existência de agentes nas redes que se encontram ao alcance do dispositivo móvel.

A motivação para o desenvolvimento da Arquitetura *Inception* para o protocolo *SMIP* advém da possibilidade de que o Agente *SMIP* instalado em uma rede móvel não precise ser o *NAT* da referida rede e, conseqüentemente, não necessite de um *IP* público para ser alcançável de outras redes com as quais o dispositivo móvel esteja interagindo para efetuar o processo de troca de redes.

O presente texto estará organizado em Trabalhos Relacionados, Proposta, Metodologia, Resultados, Conclusões / Trabalhos Futuros e Referências.

## 2. Trabalhos Relacionados

Em [Monteiro 2012] é definido o Protocolo *SMIP*: um ambiente de apoio à interconexão entre redes sem fio heterogêneas, de forma que as conexões ativas de um dispositivo móvel, em trânsito entre duas redes, não sejam perdidas, através do uso de túneis e roteamento dinâmico de pacotes.

Em [Rolim et al. 2013] é feita a implementação e a avaliação de soluções de gerência de mobilidade com o protocolo *SMIP*, com testes de desempenho demonstrando a eficiência do protocolo em relação às implementações existentes. Durante a etapa de avaliação se chegou à conclusão da necessidade de tornar o *SMIP* menos dependente da sua instalação em um *NAT*.

[Rosenberg et al. 2003] apresenta e define um protocolo chamado *Session Traversal Utilities for NAT (STUN)*, que serve como ferramenta e referência para que outros protocolos possam lidar com a travessia de dispositivos tradutores de endereços de rede (*NATs*). Ele pode ser utilizado por um ponto de acesso para determinar o IP e a porta alocados para o mesmo no *NAT*. Este é o protocolo originalmente utilizado para definir o comportamento da Arquitetura *Inception*.

Em [Peters 2006] é feita uma análise de várias implementações de *NATs*, focadas no entendimento dos mecanismos subjacentes de mapeamento de portas de comunicação e tradução de endereços de rede. Tal análise é utilizada para avaliar a viabilidade do uso do protocolo *STUN* no contexto da Arquitetura *Inception*.

## 3. Proposta

Este artigo originalmente complementa a arquitetura do Protocolo *SMIP*, definido em [Monteiro 2012], através da remoção da necessidade de *IPs* públicos nos Servidores *SMIP*. Este complemento é feito através da utilização do mapeamento de portas dinâmicas em servidores tradutores de endereços *NATs*, que operam como pontos de conexão entre a rede local e redes externas, permitindo que o *SMIP* opere de dentro da rede local para a qual ele oferece os serviços de mobilidade *IP*.

Conforme apresentado, o *SMIP* é um protocolo para gerência de mobilidade, que age combinando formas reativas e preditivas de se reduzir a latência de *handover* através da diminuição da sinalização *MIP* ao tempo necessário para que o móvel faça a requisição de montagem do túnel entre seu *Care-off Address* e o *Home Agent*, fazendo com que o *Home Agent* realize a otimização de roteamento correspondente, visando entregar os datagramas pertencentes à chamada em andamento do usuário, através da rede estrangeira.

A técnica de atravessamento de redes chamada *STUN*, conforme definida em [Rosenberg et al. 2003], apresenta uma forma de atravessamento de redes que permite que máquinas que se encontrem em uma rede privada e se comunicando com a internet através de um servidor *NAT* possam trocar dados entre si sem a necessidade de outra máquina intermediando a conexão.

### 3.1. Arquitetura Inception

A arquitetura *Inception* foi desenvolvida com o objetivo de efetuar a sincronização de portas entre clientes e servidores, quando os servidores se encontrarem atrás de um dispositivo *NAT*. Foi originalmente definida como uma alteração do funcionamento do protocolo *SMIP*, de forma a suportar atravessamento de *NAT* em situações onde não é possível

configurar o dispositivo com o Servidor *SMIP*. A arquitetura originalmente implementa o protocolo *STUN*, de [Rosenberg et al. 2003], com alterações para utilizar protocolos de mapeamento de portas, em discordância ao protocolo original, que utiliza o comportamento de gerenciamento de portas do *NAT*.

Os protocolos *UPnP* (*Universal Plug and Play*), apresentado por [Boucadair et al. 2013a], e o *NAT-PMP* (*Network Address Translation - Port Mapping Protocol*), definido em [Boucadair et al. 2013b], permitem que aplicações dentro de uma rede privada possam coordenar, junto ao *NAT*, o processo de abertura e de mapeamento de portas de forma que tais aplicações possam ser alcançáveis por clientes que se encontrem na internet. Desta forma, os protocolos oferecem uma forma padronizada e controlável de se interagir com os pontos de acesso à redes externas e à internet.

Tendo como base o uso de tais protocolos, esta proposta apresenta uma infraestrutura mista, que permite a abertura de portas dinâmicas em *NATs* que possuam implementados os protocolos *UPnP* ou *NAT-PMP*, e em seguida faça o registro das portas abertas e dos serviços oferecidos, permitindo que os clientes, através do recurso de busca de portas, possam se conectar aos serviços oferecidos, conforme o funcionamento padrão do *STUN*. Desta forma, *Inception* apresenta as seguintes entidades funcionais:

- **Servidor *Inception*** - Responsável por solicitar a abertura de portas dinâmicas ao *NAT* e fazer a configuração do Servidor *SMIP*.
- **Mapeador de Portas *NAT*** - Responsável por receber solicitações de abertura de portas e efetuar o mapeamento das mesmas para os dispositivos correspondentes.
- **Cliente *Inception*** - Responsável por iniciar uma conexão com o Site Parceiro, buscar um Servidor *Inception* compatível e fazer a configuração do Cliente *SMIP*.
- **Site Parceiro** - Responsável por armazenar as informações de portas dos Servidores *Inception* e repassar para o Cliente *Inception* para sincronização.

### 3.1.1. Servidor *Inception*

A principal tarefa do Servidor *Inception* é fazer a comunicação com o *NAT* para abertura de portas através dos protocolos *UPnP* ou *NAT-PMP*. Caso este não consiga se comunicar com o *NAT* e fazer o registro de portas, sua execução é terminada e uma mensagem de erro é disparada.

Caso haja comunicação com o *NAT* e o mapeamento de portas for efetuado, o próximo passo do Servidor *Inception* é fazer a inicialização do serviço que irá se utilizar da porta mapeada. Terminada a inicialização do serviço, o Servidor *Inception* efetua o registro da porta no Site Parceiro.

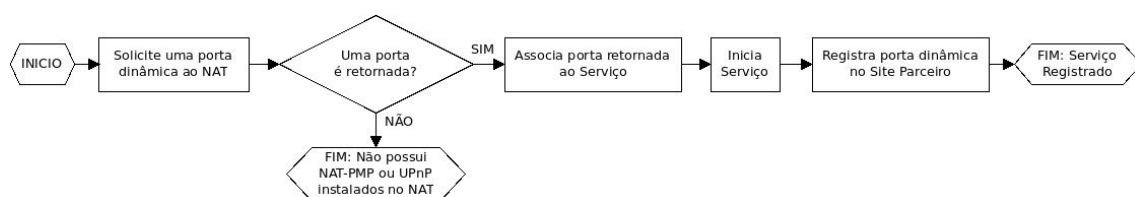


Figura 1. Fluxograma do Servidor *Inception*

### 3.1.2. Cliente Inception

O cliente *Inception* têm como objetivo fazer a sincronização do cliente com o servidor, através da busca de portas *NAT* mapeadas em um determinado *IP* que o cliente queira conectar. Para isto, o Cliente *Inception* faz uma requisição para o Site Parceiro sobre a existência de portas abertas no mesmo.

Caso nenhuma porta seja retornada, Cliente *Inception* assume que não há serviço compatível rodando no *IP* especificado e uma mensagem de erro é disparada. Caso haja porta listada, o Cliente *Inception* procede com a inicialização do cliente, utilizando a porta recebida do Site Parceiro.

À partir do ponto em que a sincronização foi efetuada e o cliente foi iniciado, o mesmo passa a funcionar independentemente da arquitetura.



Figura 2. Fluxograma do Cliente *Inception*

### 3.1.3. Site Parceiro

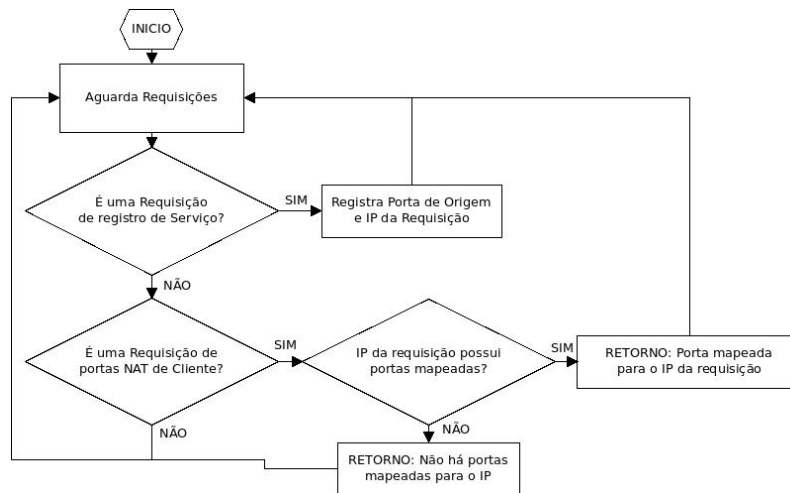
A arquitetura *Inception* necessita de um servidor público para funcionar, e este servidor precisa estar visível tanto para o cliente quanto para o servidor. É através de convenção por configuração, que o Site Parceiro será utilizado. Então, durante o processo de inicialização do Servidor *Inception*, o mesmo irá registrar a porta que foi aberta no *NAT* e, quando um Cliente *Inception* solicitar uma busca pelo *IP* do *NAT* da sua rede, o Site Parceiro irá retornar a porta *NAT* registrada, permitindo que cliente e servidor possam se comunicar.

Ao contrário do Cliente e do Servidor *Inception*, o Site Parceiro sempre escuta por novas requisições.

Desta forma, a arquitetura *Inception* funciona como um intermediário entre o Servidor e os Clientes *SMIP*, fazendo a atribuição automática das portas para ambas as partes e permitindo que o *SMIP* funcione sem que haja alteração significativa no seu modo de funcionamento.

## 4. Metodologia

Inicialmente, foi proposta a alteração do protocolo *SMIP* de forma que o mesmo incorporasse funcionalidades de abertura de portas através da técnica *STUN*, definida em [Rosenberg et al. 2003]. Desta forma, foi criada uma entidade adicional que ficaria responsável por efetuar o sincronismo entre cliente e servidor, através da abertura e manutenção da porta através de pacotes de manutenção da conexão e da publicação de tais portas, tornando as mesmas detectáveis para os clientes que necessitem de conexão.



**Figura 3. Fluxograma do Site Parceiro**

O processo consiste em utilizar o Site Parceiro para que o servidor efetue uma conexão abrindo, deliberadamente, uma porta dinâmica no *NAT* que conseqüentemente é registrada no Site Parceiro para ser entregue a um cliente. À partir de então o cliente, de posse dessa informação, inicia uma conexão com a mesma, alcançando o serviço oculto através da porta mapeada.

No entanto, durante o processo de implementação da solução *STUN* foi descoberto que determinadas configurações de *NAT* poderiam, de fato, impedir o funcionamento pleno da solução. Tais configurações se referem ao comportamento do *NAT* em resposta à conexão de diferentes hosts a uma mesma porta. Segundo [Peters 2006], este comportamento é esperado em *NATs* do tipo Simétrico.

Durante a fase de avaliação da viabilidade da nova solução, foi proposto que a arquitetura de abertura de portas fosse separada da arquitetura do protocolo *SMIP*, tornando-se assim um intermediário entre cliente e servidor. Essa separação, que visava a manutenção de um código mais simples, também permitiu que a arquitetura pudesse ser utilizada não somente para o *SMIP* em questão, mas qualquer modelo cliente-servidor, cuja entidade servidora se encontrasse atrás de uma rede mascarada.

Verificada a viabilidade da nova solução, foi iniciado o processo de desenvolvimento da mesma, utilizando-se como base a linguagem de programação *Python* usando *Sockets* Assíncronos no Cliente e no Servidor e *Python* com *Flask* no Site Parceiro.

## 5. Resultados

Através do desenvolvimento e dos testes realizados, foi demonstrada a viabilidade do uso da arquitetura *Inception* como mediadora para conexões entre clientes e servidores em uma topologia de redes mascaradas sem a necessidade se alterar o funcionamento interno dos mesmos. Isto foi possível através do uso de protocolos de mapeamento de portas, que permitiram o controle do processo de *Hole Punching* executado pelo servidor atrás da rede mascarada.

É importante ressaltar que o uso da Arquitetura *Inception* em conjunto com o *SMIP* só ocorre nos contextos de rede em que não for possível instalar o Servidor *SMIP* no *NAT*, devido restrições de acesso ou segurança vigentes. Nos casos em que o *SMIP* não

pode ser executado no *NAT*, situação para a qual o protocolo *SMIP* não foi originalmente desenvolvido, o uso da arquitetura *Inception* deve permitir a execução do referido serviço, sem perdas aparentes no processo. Esta solução se mostrou eficiente neste cenário.

Outro resultado encontrado foi a possibilidade do uso da arquitetura com qualquer topologia de cliente-servidor, nos parâmetros topológicos abordados no decorrer do artigo, através da sincronização que é efetuada entre Servidor *Inception*, Cliente *Inception* e Site Parceiro antes da inicialização e posterior estabelecimento da conexão entre os serviços através do *NAT*.

Através destes resultados, fica constatado que a arquitetura *Inception* é capaz de inferir a rede em busca da possibilidade de ser efetuado o atravessamento do *NAT* através do teste de compatibilidade do mesmo com os protocolos de abertura de portas dinâmicas. Também fica demonstrado que os componentes do protocolo *SMIP* podem funcionar independente das estruturas de *NAT* subjacentes que existam entre estes.

## 6. Conclusões / Trabalhos Futuros

A arquitetura provou-se viável em tornar visíveis serviços instalados atrás de roteadores que utilizam *NAT*. Usada em conjunto com o *SMIP*, a arquitetura permite ao mesmo a flexibilidade de não ser necessária a instalação no próprio *NAT*, algo que nem sempre é possível, dadas as condições administrativas de cada rede envolvida no processo.

Eliminando-se a necessidade de se ter um *IP* público na máquina rodando o *SMIP* o mesmo pode, então, ser executado em qualquer máquina dentro da rede privada onde se encontra o móvel e que irá fazer a transição de redes usando o protocolo de mobilidade.

Trabalhos futuros podem abordar o impacto que o uso que a arquitetura *Inception* adiciona ao processo de *handover* do protocolo *SMIP*, medindo o nível de degradação da latência para a sincronização do cliente no processo de troca de redes.

Outros trabalhos relacionados podem fazer uso da arquitetura para criar uma rede de serviços distribuídos através do Site Parceiro, que passaria a ser tratado como um Serviço de Diretório para serviços que se encontram em redes mascaradas, como Redes Privadas Virtuais (*VPNs*).

## Referências

- Boucadair, M., Penno, R., and Wing, D. (2013a). Universal plug and play (upnp) internet gateway device-port control protocol interworking function (igd-pcp iwf).
- Boucadair, M., Selkirk, P., Penno, R., Wing, D., and Cheshire, S. (2013b). Port control protocol (pcp).
- Monteiro, C. d. C. (2012). Um ambiente para apoio à integração de redes sem fio heterogêneas.
- Peters, H. (2006). Analysis of nat approaches and explicit signaling for nat traversal. *April2006*.
- Rolim, E. N. d. S., Souza, J. A., Meira, L. H., Oliveira, R. d. L., and Monteiro, C. d. C. (2013). Implementação e avaliação de soluções para gerência de mobilidade ip com o protocolo smip.
- Rosenberg, J., Mahy, R., Huitema, C., and Weinberger, J. (2003). Stun-simple traversal of user datagram protocol (udp) through network address translators (nats).

# RoboCar: protótipo de robô móvel autônomo

Luan Teylo G. Lima<sup>1</sup>, Danilo Junior M. Londero<sup>1</sup>, Mathias F. Mantelli<sup>1</sup>, Livia Lopes Azevedo<sup>1</sup>

<sup>1</sup>Ciência da Computação – Universidade Federal de Mato Grosso (UFMT)  
78600-000 – Barra do Garças – MT – Brazil

{luanteylo,danilolondero,mathiasfassini}@gmail.com, livia@ufmt.br

**Abstract.** *In this paper, we would like to report the development of an autonomous mobile robot, RoboCar, which uses the LEGO MINDSTORMS NXT kit. Our goal is to discuss the physical structure of the prototype, the distribution of sensor and actuators, the development of a control system, and the choice of the architecture that underpins programming. The results in a simulated scenario suggest that the RoboCar manages to perform the route autonomously, respecting intersections and avoiding collisions.*

**Resumo.** *Este trabalho apresenta o desenvolvimento de um robô móvel autônomo, denominado RoboCar, fazendo uso do kit LEGO MINDSTORMS NXT. Nele será discutido a estrutura física do protótipo, a distribuição dos sensores e atuadores, a elaboração de um sistema de controle e a escolha da arquitetura que sustenta sua programação. Os resultados obtidos mostram que o RoboCar foi capaz de realizar o percurso de forma autônoma no cenário simulado, respeitando cruzamentos e evitando colisões.*

## 1. Introdução

Chama-se de veículo móvel autônomo qualquer veículo terrestre que seja capaz de, sem a intervenção direta de humanos, transportar pessoas e objetos aos destinos desejados. Porém, não basta o veículo sair de sua origem e chegar até seu destino, ele deve, ainda, garantir um transporte seguro e confiável. Essas ações são possíveis de serem realizadas graças a uma interação muito forte entre sensores, sistemas de controle e atuadores que permitem a esses mecanismos escolher a melhor opção de decisão e ação baseado no ambiente no qual está inserido.

Segundo [Honório et al. 2010; Fernandes et al. 2010; Atreya et al. 2006], diversos grupos, no Brasil e no exterior, têm trabalhado na pesquisa e desenvolvimento de automóveis autônomos ou semiautônomos. Segundo [Lima e Pereira 2011] um dos objetivos destes grupos é desenvolver tecnologias que possam ser integradas a veículos comerciais, dotando-os de sistemas que aumentem o conforto e a segurança de seus motoristas e passageiros.

Nesse contexto, de acordo com [Araújo e Librantz 2006], os estudos e o desenvolvimento de projetos de robôs (ou agentes) autônomos podem estar ligados a diversas aplicações práticas e têm suscitado grandes desafios, visto que as dificuldades se multiplicam à medida que o ambiente torna-se mais imprevisível e diversificado. Nessas condições o robô deve ser capaz de estimar uma posição no ambiente, reconhecer obstáculos, adquirir e utilizar conhecimento sobre o ambiente de modo a responder em tempo real as situações que possam ocorrer.

Desenvolver uma arquitetura de controle que possibilite um robô móvel autônomo operar em um ambiente, de modo estável e robusto, tem sido o desafio da área de automação. Além disso, essa arquitetura deve possibilitar flexibilidade e modularização para lidar com vários tipos de tarefas, bem como a possibilidade de alteração ou adição de novos hardwares. Para validar o desenvolvimento de uma

arquitetura de hardware e software capaz de permitir um robô se locomover de um ponto A para um ponto B, evitando obstáculos e percebendo eventuais mudanças nesse ambiente, foi montado um protótipo denominado **RoboCar**.

## 2. Arquiteturas

Há na literatura várias propostas de arquiteturas, porém não existe um paradigma definitivo que atenda todas as funcionalidades requeridas por um sistema robótico autônomo. Dowling [apud. Junior2006] descreve uma arquitetura para robôs como a solução de software e hardware usada para desenvolver o sistema de controle, incluindo a forma como os componentes da arquitetura comunicam entre si. Uma das formas de se classificar uma arquitetura para robôs móveis está relacionada ao uso de deliberação e reatividade dentro do sistema de controle. Segundo este critério, as arquiteturas podem ser divididas em: deliberativas, reativas e híbridas [Junior 2006; Silva 2009; Ribeiro, Costa e Romeiro 2001; Heinen 2002]. Esses autores classificam as mesmas como seguem:

### 2.1. Arquiteturas deliberativas

Em arquiteturas deliberativas, o agente (robô) é capaz de planejar ações futuras, possui conhecimento interno do ambiente e pode tomar decisões com base nesse conhecimento. Segundo [Junior 2006], a abordagem deliberativa transcreve o processo de planejamento e tomada de decisão do homem para que o robô execute determinada atividade. Nessa arquitetura as atividades são previamente planejadas e as ações ocorrem de forma sequencial e organizada. O robô estabelece o mapa, planeja todas as ações necessárias para atingir o objetivo e, na sequência, age para atingir o objetivo. Depois de realizar toda a sequência ele refaz o planejamento e começa o ciclo novamente.

Uma das principais vantagens desse modelo é a precisão oferecida por esse, com relação à localização do robô, bem como a execução de tarefas mais complexas que exigem planejamento. Entretanto, essa também se torna a pior desvantagem do modelo, uma vez que o custo computacional para a atualização do ambiente e realização das atividades é muito maior. Esse modelo não atua com robustez em sistemas dinâmicos, por exemplo, com o surgimento de obstáculos. Como exemplo de arquiteturas deliberativas pode-se citar NHC (Controlador Hierárquico) [Meystel 1991] e RCS (NIST Sistema de controle de tempo real) desenvolvida por [Albus e Meystel 1996].

### 2.2. Arquiteturas reativas

Arquiteturas reativas são baseadas em comportamentos. Nessas, dado uma condição sensorial, o robô então executa um ação pré-determinada. “O principal objetivo das arquiteturas reativas é possibilitar a implementação de sistemas de controle que respondam rapidamente a uma variedade de ocorrências ou situações no ambiente, fazendo com que robôs operem em ambientes altamente dinâmicos”, [Junior 2006]. Uma das vantagens dessa arquitetura é que, em geral, a velocidade de processamento e resposta é alta, devida à simplicidade no tratamento das informações e devido à maneira direta pela a qual a percepção, ou estímulo, está associado com uma ação, ou resposta [Junior 2006].

Por outro lado, a desvantagem se dá no fato de inexistência de planejamento perante as ações, uma vez que o comportamento se dá por reflexão, não permitindo a execução de atividades mais complexas. Como exemplo de arquiteturas reativas pode-se citar a Arquitetura de Subsunção [Brooks1986] e Motor Schema [Arkin 1998].

### 2.3. Arquiteturas híbridas

Arquiteturas híbridas, como sugere o nome, são feitas em camadas que integram componentes deliberativos com componentes reativos. Após serem planejadas, por meio dos componentes deliberativos, as ações são executadas utilizando a reatividade. Isso permite que o agente responda rapidamente a mudanças do ambiente. Segundo [Junior 2006], a arquitetura híbrida procura ser adequada para solucionar problemas complexos. A deliberação é responsável por atingir os objetivos de maneira ótima e eficiente. Já a reatividade garante que as operações sejam realizadas em ambientes dinâmicos que exigem rapidez na resposta.

A arquitetura híbrida, conforme descrito por [Murphy 2000], geralmente apresenta um conjunto básico de componentes: Sequenciador; Gerenciador de recursos; Cartógrafo; Planejador de missão; Monitoração de desempenho e resolução de problemas. A grande vantagem dessa arquitetura reside na existência de planejamento aliado a percepção e ação, possibilitando que as ações de planejamento, que são mais lentas, possam ser executadas paralelamente as ações de percepção e ação. A desvantagem dessa arquitetura refere-se, também, ao gasto computacional relacionada à utilização atualizada do ambiente. Além do sincronismo inerente à execução em paralelo das ações de planejar, perceber e agir. Como exemplo de arquiteturas híbridas pode-se citar AuRA [Arkin 1986] e SFX [Murphyn e Arkin1992].

## 3. O protótipo do RoboCar

O protótipo do robô móvel autônomo – RoboCar é descrito na sequência. Sendo este desenvolvido em duas etapas: arquitetura e estrutura.

### 3.1. Arquitetura de software

Neste trabalho, optou-se por uma arquitetura híbrida, pois o RoboCar precisaria estar apto a trabalhar em ambientes dinâmicos. A arquitetura implementada foi baseada na arquitetura AuRA [Arkin 1986]. Trata-se de uma arquitetura modular disposta em camadas, Figura 1. Cada módulo resolve um subproblema específico de seu domínio, tendo como entrada, normalmente, o resultado de um módulo anterior. Os módulos são dispostos da seguinte forma:

- **Mapa:** Apresenta uma informação estática do ambiente, sendo que essa deve ser passada para o robô antes do início do programa. Para representar o ambiente, utilizou-se uma matriz de ocupação binária. Essa técnica simplifica a construção e a representação do mapa, além de facilitar a determinação de caminhos. No entanto, por se tratar de uma matriz, há um alto consumo de memória, o que a torna imprópria para ambientes complexos. Porém, nesse trabalho, utilizou-se um ambiente controlado e simples o que viabilizou o uso da matriz de ocupação binária.
- **Planejador de Rotas:** Utiliza as informações de posição inicial e posição final para determinar a rota que será executada pelo robô. O Planejador de Rotas aplica o algoritmo de Busca em Profundidade no Mapa, e então guarda o caminho de forma sequencial em um vetor que será então passado para o Piloto.
- **Piloto:** Determina quais movimentos serão realizados para que o objetivo seja cumprido. O Piloto cruza as informações da direção atual do robô (captada por um sensor bússola), sua posição em relação ao mapa e próxima posição a ser alcançada; infere qual movimento deve ser feito para chegar à próxima célula. A saída desse módulo é um comando para a camada reativa.
- **Controlador de Movimento:** Recebe um comando que determina qual movimento deverá ser feito e traduz essa informação para os atuadores. Ao executar uma ação, o Controlador de Movimento deve manter a integridade do robô, para isso é executado,

em paralelo, um módulo que determina a chance de colisão com base na leitura do sensor ultrassônico. Quando a chance de colisão é alta, o módulo faz uma intervenção e determina que um movimento de evasão deve ser executado.



Figura 1. Arquitetura de controle

### 3.2. Estrutura do RoboCar

A estrutura do RoboCar foi desenvolvida usando o kit LEGO MINDSTORMS, versão 1.0, um kit de robô programável voltado para a educação tecnológica. O protótipo, foi montado sobre uma base de quatro rodas, duas na dianteira e duas na traseira, onde está a tração, proporcionada por dois servo-motores, figura 2. Essa configuração ofereceu um melhor desempenho nas curvas, situação em que as duas rodas traseiras giram em velocidades diferentes. As rodas dianteiras possibilitam ao RoboCar virar tanto para a direita quanto para a esquerda, utilizando-se um servo motor para essa ação.

A percepção do robô foi realizada com a utilização de três sensores: o ultrassônico, responsável por evitar colisões entre o robô e as paredes; a bússola, responsável por dar a direção e possibilitar ao robô determinar o sentido; e o sensor de luz, responsável por identificar o fim das ruas e o início dos cruzamentos, por meio de marcações artificiais no ambiente.

O sensor ultrassônico foi posicionado na lateral direita do robô para permitir a este evitar colisões e manter o movimento paralelo às paredes, ou seja, caso se aproxime demais da parede seu movimento é de se afastar e em caso contrário, seu movimento é de aproximação. O sensor de luz foi posicionado na dianteira do robô para permitir a este identificar as marcações do ambiente, isto é, o fim de rua e início de cruzamento. A bússola também foi posicionada na dianteira do robô, melhor posição considerada, pois ficava livre de interferência magnética causada pelos outros componentes.

O Brick NXT (tijolo) foi fixado sobre a estrutura que liga os eixos. O tijolo é o cérebro do robô e responsável por executar os programas, analisando as entradas (sensores) e coordenando as saídas (atuadores), nele são conectados todos os sensores e todos os atuadores. A figura 2 apresenta a disposição dos sensores: 1) Sensor ultrassônico; 2) Bússola; 3) Sensor de luz.

O tijolo possui como micro controlador principal um Atmel® 32-bit ARM® com 48 MHz, 64 KB RAM e 256 KB FLASH, um micro controlador secundário Atmel® 8-bit AVR com 8 MHz, 512 Byte RAM e 4 KB FLASH. O sensor ultrassônico é capaz de detectar objetos entre 0 e 255 centímetros de distância com precisão de 1 centímetro. O sensor de luz captura a luminosidade do ambiente, numa leitura de 0 a 1023, correspondendo a sua intensidade numa variação de mais escuro para o mais claro. A bússola possui um sensor Honeywell HMC5883L com precisão de 1° a 2°.



Figura 2. Protótipo RoboCar

#### 4. Resultados alcançados

Este projeto deu ênfase à construção de um veículo – RoboCar, capaz de se locomover de forma autônoma em ambientes que simulassem situações reais. O cenário utilizado para testes representava as ruas de uma cidade, onde o RoboCar, saindo de um ponto específico, recebia o ponto de destino, calculava uma rota, e esta deveria ser percorrida de modo autônomo, respeitando cruzamentos e evitando colisões. Para realizar tal tarefa, o RoboCar contou com uma arquitetura de controle que foi desenvolvida para possibilitar ao mesmo operar no ambiente determinado, utilizando seus recursos físicos e computacionais de modo mais robusto e estável.

Com exceção da bússola, o protótipo foi montado utilizando-se apenas peças e sensores presentes no kit LEGO MINDSTORMS NXT. O sistema de controle permitiu a movimentação autônoma do robô dentro do ambiente. A interação entre as camadas, deliberativa e reativa, fez com que o protótipo pudesse determinar um caminho a ser seguido e, na sequência, executar as ações necessárias para cumprir o percurso.

Durante a bateria de testes foram alcançados os principais objetivos da proposta. Com relação à arquitetura de controle percebeu-se que: na parte deliberativa, o sistema demonstrou-se eficiente para determinar a rota a ser seguida, mesmo em situações onde o protótipo era colocado de forma não convencional na posição inicial, isto é, quando este não está de frente ao caminho a ser seguido. Nessa condição, o sistema foi capaz de manobrar para normalizar a posição e então seguir a trajetória. A parte reativa, com base nas leituras dos sensores, respondeu prontamente com as ações a serem executadas como: parar nos cruzamentos e manobrar o veículo evitando colisões.

Com a elaboração do protótipo foi possível perceber uma variedade de dificuldades relacionadas tanto à escolha da arquitetura quanto a estrutura física. Uma delas está relacionada ao fato do NXT ter um baixo poder de processamento, o que impôs a necessidade de simplificações ao sistema para melhorar o seu desempenho.

Outra preocupação foi a falta de exatidão do sensor ultrassônico disponível no kit. Este, por se tratar de um dispositivo simples, projetado para trabalhar em um meio que não sofre grandes alterações, não conta com uma forma de calibragem, diferentemente dos sensores de som e de luz. Assim sendo, fez-se necessário desenvolver uma função para realizar a correção das informações de distâncias informadas pelo sensor antes de se trabalhar com elas.

As correções foram realizadas por meio de testes com diferentes medições, entre 1 cm a 22 cm, verificando as diferenças entre os valores lidos pelo sensor e o valor real, a fim de se ter uma média da margem de erro. Essa margem de erro foi calculada tanto com o sensor parado quanto em movimento, e informações como temperatura e umidade do ambiente não foram consideradas. Com os cálculos, foi constatado um erro de 3 centímetros para o sensor utilizado. A função desenvolvida trabalhava não apenas refinando as

medições com base na margem de erro, mas também calculando uma média ponderada dos últimos valores lidos e exercia, também, a tarefa de filtragem, desconsiderando valores não condizentes com o ambiente.

Como atividades futuras pretende-se incorporar ao kit LEGO outros dispositivos de hardware e peças de montagem (engrenagens, sistemas de direção, diferencial, entre outros) que possibilite um protótipo mais fiel a um veículo real, melhorando a interação do protótipo com o ambiente. Com relação a estrutura de controle, pretende-se capacitar o robô de modo que este possa oferecer mais funcionalidades como: estacionar, desviar de obstáculos moveis, dar ré. Pretende-se também aperfeiçoar a função existente de recalculer a rota na ocorrência de obstáculos.

## 7. Referências

- Araújo, S.A. and Librantz, A.F.H. (2006) “Navegação Autônoma de Robôs”, In: *Exacta*, São Paulo, v. 4, n. especial, p. 81-83.
- Arkin, R. C. (1986) “Path Planning for a Vision-Based Autonomous Robot”, In: Proceedings of the SPIE Conference on Mobile Robots, Cambridge, MA.
- Albus, J.S. and Meystel, A. (1996) "A Reference Model Architecture for Design and Implementation of Intelligent Control in Large and Complex Systems", In: *International Journal of Intelligent Control and Systems*, Vol. 1, No. 1, pg. 15-30.
- Atreya, A. R., Cattle, B. C., Collins, B. M., Es-Senburg, B., Franken, G. H., Saxe, A. M., Schiffres, S. N. and Kornhauser, A. L. (2006) “Prospect eleven: Princeton university’s entry in the 2005 DARPA Grand Challenge”, In: *Journal of Field Robotics*, v.23, n. 9, p. 745–753.
- Brooks, R. A. (1986) “A Robust Layered Control System for a Mobile Robot”, In: *IEEE Journal Of Robotic and Automation*.
- Fernandes, L. C., Dias, M. A., Osorio, F. and Wolf, D. F. (2010) “A driving assistance system for navigation in urban environments”, In: Proceedings of the Ibero-American Conference on Artificial Intelligence.
- Junior, V. G.(2006) “Arquitetura Híbrida para Robôs Móveis Baseada em Funções de Navegação com Interação Humana”, São Paulo.
- Heinen, F. J. (2002) “Sistema de Controle Híbrido para Robôs Móveis Autônomos”, São Leopoldo-SP.
- Honorio, L. M., Vermaas, L. L. G., Goncalves, L. M. and Vidigal, M. (2010) “Uma metodologia para aprendizado supervisionado aplicada em veículos inteligentes”, In: XVIII Congresso Brasileiro de Automática, pp. 1028–1035.
- Meystel, A. (1991) “Nested Hierarchical Control”, Philadelphia.
- Murphy, R. and Arkin, R. (1992) “Sfx: An architecture for action-oriented sensor fusion”, In: 1992 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 225-250.
- Ribeiro, C. H. C., Costa, A. H.R. and Romero, R. A. F. (2001) “Robôs Móveis Inteligentes: Princípio e Técnicas”.
- Santos, K. R. da S. (2009) “Sistema de Navegação Autônoma para Robôs Móveis Baseado em Arquitetura Híbrida: Teoria e Aplicação”, Itajubá-MG.

# Sistema para registro e manipulação de dados de precipitação de chuva do projeto TRMM

Gracyeli S. Souza<sup>1</sup>, Ivairton M. Santos<sup>2</sup>, Peter Zeilhofer<sup>1</sup>

<sup>1</sup> Programa de Pós-Graduação em Física Ambiental – Universidade Federal de Mato Grosso (UFMT) – Cuiabá, MT– Brasil

<sup>2</sup> Instituto de Ciências Exatas e Da Terra – Campus Universitário do Araguaia – Universidade Federal de Mato Grosso (UFMT) – Barra do Garças, MT – Brasil  
gracyeli@gmail.com, ivairton@ufmt.br, zeilhoferpeter@gmail.com

**Abstract.** *The Tropical Rainfall Measuring Mission (TRMM) is an important source of rainfall data. The formats, however, in which the data are being made available, hamper its usage, as they include large amounts of unstructured and non-normalized data after download. This paper presents a web-based environment that stores and manipulates data of the 3B42 TRMM product, facilitating its access in an optimized response time through a graphical user interface. The system replicates recorded data, guaranteeing its integrity and consistency, allows spatio-temporal queries through a map server connected to a geographical database by common web browsers.*

**Resumo.** *O projeto Tropical Rainfall Measuring Mission (TRMM) é uma importante fonte de dados de precipitação. No entanto os formatos em que os dados são disponibilizados dificultam sua utilização, por constituírem um grande volume não estruturado e normalizado após o download. Este trabalho apresenta um sistema em ambiente web, que armazena e manipula dados do conjunto 3B42 da TRMM, facilitando seu acesso em tempo de resposta otimizado a partir de uma interface gráfica. O sistema replica os dados registrados, garantindo sua integridade e consistência, permite buscas espaço-temporais a partir de um servidor de mapas conectado a um banco de dados geográfico em ambiente web.*

## 1. Introdução

Sistemas para monitoramento ambiental baseados em técnicas de sensoriamento remoto, de modo geral, geram grandes volumes de dados. Esses conjuntos, armazenados em Sistemas Gerenciadores de Banco de Dados (SGBD) demandam técnicas computacionais cada vez mais avançadas para sua extração, sintetização e mineração.

Potenciais usuários destes conjuntos de dados geralmente não dispõem de conhecimentos aprofundados em computação, demandando por ferramentas interativas e de manuseio intuitivo para a utilização efetiva destes dados. Sendo estes conjuntos de dados referenciados no tempo e espaço, é preciso ferramentas que permitam o acesso aos dados de forma eficiente. Neste contexto, o Sistema de Informação Geográfica (SIG) em ambiente *web*, desenvolvido neste trabalho, visa auxiliar no acesso aos dados de precipitação provenientes do projeto Tropical Rainfall Measuring Mission (TRMM), disponibilizado pela *National Aeronautics and Space Administration* (NASA).

## 2. TRMM

Estudos da precipitação e sua distribuição no espaço e tempo são importantes para o entendimento e gerenciamento sustentável dos recursos naturais, pois afetam diretamente múltiplos setores da sociedade e do meio ambiente. São várias as consequências decorrentes do excesso ou escassez das chuvas, que pode ocasionar desastres ambientais, problemas de abastecimento hídrico, sanitário, afetando diversos setores, como por exemplo, a produção agrícola (PERETTI e NUMMER, 2012).

Com o objetivo de monitorar a precipitação nos trópicos e servir como fonte operacional de dados de chuva, a NASA em parceria com a *Japan Aerospace Exploration Agency* (JAXA) lançou o satélite do projeto TRMM no dia 27 de novembro de 1997 (NASA, 2014). O satélite foi desenvolvido de maneira que pudesse medir os índices de precipitação das chuvas nas latitudes entre 35°N e 35°S (KUMMEROW *et al.*, 2000). O registro de dados ocorre para a maior parte do globo terrestre (exceto os polos), incluindo grandes áreas do oceano onde não há equipamentos instalados para medição de precipitação (MACEDO, 2013). Portanto, os sensores disponíveis no satélite permitem o monitoramento da precipitação inclusive de regiões de difícil acesso ou que não possuem estações de monitoramento convencionais.

Este estudo utiliza dados do produto 3B42, provenientes do projeto TRMM, gerados desde 1998, em resolução temporal de três horas e resolução espacial de 0.25° (NÓBREGA, *et al.*, 2008). Os conjuntos de dados deste produto somam terabytes de volume de dados, que são disponibilizados em dois tipos de formatos binário, *Hierarchical Data Format* (HDF) ou *Network Common Data Form* (NetCDF). De modo geral, estes formatos dificultam o seu acesso, manipulação e análise.

A complexidade inerente à manipulação dos dados disponibilizados pelo projeto TRMM está presente nos processos de aquisição, processamento e seleção dos dados de precipitação. Sua utilização requer o emprego de técnicas que possibilitam a integração com outros tipos de dados espaciais (limites de bacias, eco-regiões etc.). Dessa forma, este trabalho propõe uma ferramenta computacional que auxilie o pesquisador nesse processo de manipulação dos dados oriundos do projeto TRMM e integração com outros dados geográficos, por meio de serviços *web*, utilizando recursos de Banco de Dados Geográfico (BDG) e servidor de mapas, permitindo a ele concentrar-se apenas na manipulação e análise dos dados de precipitação.

## 3. Desenvolvimento do Sistema

O sistema aqui apresentado utiliza arquivos no formato NetCDF, onde os dados são dispostos logicamente em uma matriz, disponíveis para *download* pela NASA<sup>1</sup>.

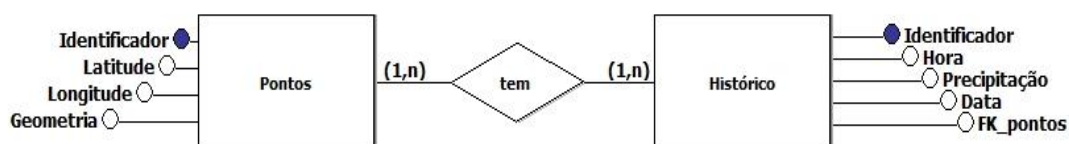
O desenvolvimento foi dividido em quatro etapas: modelagem da base de dados, criação da base de dados com suporte geográfico, configuração do servidor de mapas e desenvolvimento da interface com o usuário.

Os dados de precipitação obtidos do projeto TRMM não estão estruturados segundo as normas de normalização em banco de dados. Assim, a primeira etapa do

---

<sup>1</sup> Disponível em: <[http://disc2.nascom.nasa.gov/dods/3B42\\_V7\\_daily/](http://disc2.nascom.nasa.gov/dods/3B42_V7_daily/)>.

desenvolvimento do SIG foi a modelagem da base de dados (com informação geográfica). O Diagrama Entidade-Relacionamento mostra a entidade Pontos da grade TRMM e sua relação com o Histórico do produto 3B42 (Figura 1).



**Figura 1. Diagrama Entidade-Relacionamento do SIG web**

O processamento do conjunto de dados de precipitação nesta etapa garante a normalidade dos dados, armazenamento consistente, eficiência no acesso e evita redundância dos dados.

A segunda etapa consiste na construção da base de dados geográfica. Para armazenar os dados utilizou-se um BDG, vinculado a um servidor de mapas para facilitar e ampliar a disponibilidade dos dados via HTTP. O SGBD utilizado foi o PostgreSQL na versão 9.1, que possui código aberto, confiabilidade e integridade na transação de dados. A sua junção com a extensão espacial PostGis, que adiciona suporte a objetos geográficos, possibilita que o PostgreSQL seja utilizado como um banco de dados espacial.

Os dados de precipitação do projeto TRMM foram importados para o BDG, empregando as funções disponíveis pela extensão PostGis. Na prática, este processo consiste de um conjunto de ações que engloba a extração dos dados do arquivo fornecido pelo projeto TRMM, o armazenamento desses dados em um banco de dados intermediário, criação de um arquivo no padrão CSV<sup>2</sup>, com a informação da latitude e longitude de cada ponto, a geração de um arquivo *shapefile*<sup>3</sup> a partir do arquivo CSV, e finalmente a importação do *shapefile* para o BDG. Ao importar o *shapefile*, o PostGis cria automaticamente um atributo que contém a especificação geométrica do objeto e atribui a ele uma coordenada espacial.

O SIG *web* desenvolvido utiliza a Internet como o meio para disseminação dos dados de precipitação, para executar análises espaciais e oferecer serviços. Assim a terceira etapa do desenvolvimento do sistema abordou a configuração do servidor de mapas, que utiliza a mesma arquitetura dos sistemas para *web*, baseado no conceito cliente-servidor.

O navegador é o cliente que “conversa” com o servidor de *web*, este requisita os dados do servidor de mapas que interpreta os dados que foram solicitados, buscando as informações requeridas pelo cliente no BDG, por meio das especificações da *Open Gis Consortium* (OGC<sup>4</sup>) e então produz uma saída e esta é retornada ao cliente.

<sup>2</sup> *Comma-Separated Values* (CSV) é um formato de arquivo que armazena dados tabelados, separados por um delimitador, geralmente vírgulas e a quebra de linha para separar os valores.

<sup>3</sup> Formato de arquivo contendo dados geoespaciais vetoriais. Descreve geometria como pontos, linhas e polígonos, além dos dados relacionados aos componentes geométricos.

<sup>4</sup> Disponível em: <<http://www.opengeospatial.org/>>

O servidor de mapas facilita a disponibilidade dos dados armazenados no BDG, especialmente por proporcionar interoperabilidade entre sistemas que demandam dados manipulados, por meio de uma interface de acesso centralizada. Neste trabalho utilizou-se o servidor de mapas GeoServer, escrito em Java, que segue os padrões do OGC e possui código aberto.

Na última etapa foi desenvolvida a interface *web* do sistema, por meio da qual o usuário pode ter acesso aos dados e a visualização das informações geográficas (mapas). Para construir a interface do SIG, as páginas *web* foram desenvolvidas utilizando as linguagens HTML, PHP e JavaScript, juntamente com a API OpenLayers na versão 2.12.

O OpenLayers é responsável por renderizar o mapa no navegador, comunicando com o servidor de mapas GeoServer por meio de requisições HTTP. A utilização do OpenLayers na elaboração do sistema proporcionou facilidade e agilidade no desenvolvimento, possibilitando entre outras funcionalidades a manipulação do mapa e a seleção dos pontos de precipitação de maneira simples, proporcionando ao usuário um SIG com interface intuitiva.

A junção de todas as ferramentas e tecnologias descritas permitiu o desenvolvimento de um sistema que exibe os mapas, busque por um conjunto de dados, visualize informação dos pontos, exiba relatórios das consultas realizadas e extraia os dados de uma busca, possibilitando a exportação dos dados para um arquivo CSV.

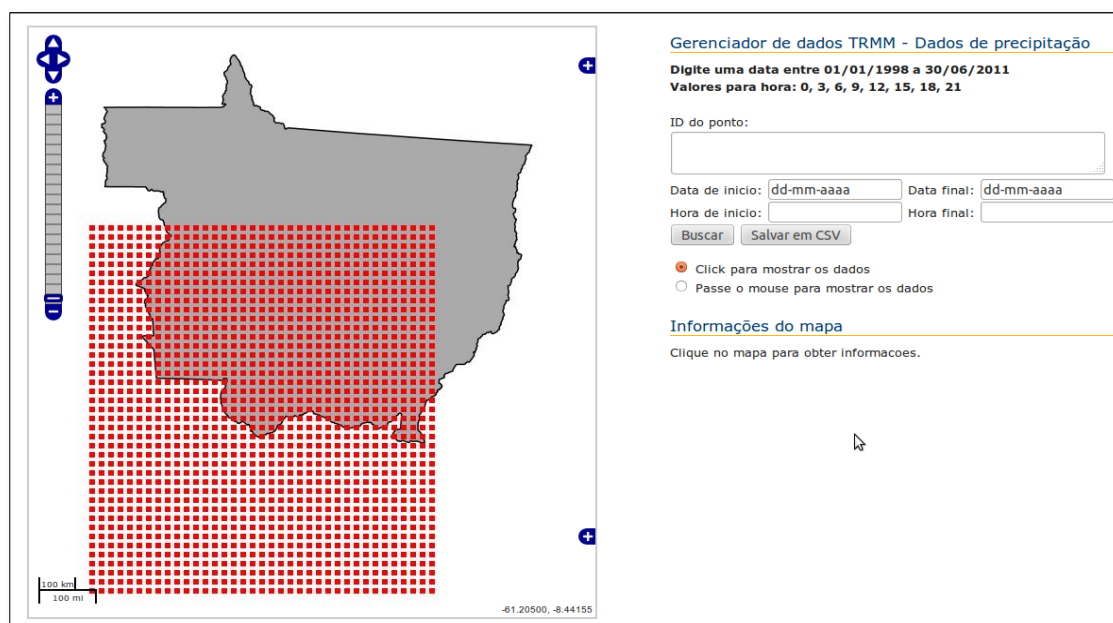
#### 4. Resultados

A aplicação desenvolvida disponibiliza ao usuário quatro funcionalidades principais (Quadro 1).

**Quadro 1. Funcionalidades do sistema.**

|                            |   |
|----------------------------|---|
| <b>Visualizar Mapa</b>     | Um mapa é apresentado ao usuário com a representação dos pontos disponíveis para consulta do histórico de precipitação.   |
| <b>Informações do Mapa</b> | Ao clicar em um ponto no mapa, são apresentados ao usuário os dados de latitude e longitude do ponto selecionado e também as informações gerais a respeito do ponto.  |
| <b>Buscar Dados</b>        | O usuário seleciona o(s) ponto(s) e define o período temporal que deseja buscar por dados. O sistema recebe a solicitação e efetua a pesquisa no banco de dados, apresentando o conjunto de dados encontrados ordenados pelos pontos. |
| <b>Guardar Busca</b>       | É fornecido ao usuário a opção de buscar dados e armazená-los em um arquivo de saída. Assim, após a busca no banco de dados, os dados encontrados são registrados em um documento.  |

No momento do acesso ao sistema é apresentada a interface (“Visualizar Mapa”), exposta na Figura 2, representando um mapa do limite do estado e um recorte da grade TRMM das regiões sul e central de MT, sobre os quais é possível aplicar efeitos de zoom e movê-lo para visualizar pontos específicos. Ao lado direito é ainda exposta informações sobre a série temporal disponível no momento (período entre 01 de janeiro de 1998 a 30 de junho de 2011).



**Figura 2 – Interface do SIG desenvolvido.**

A visualização dos dados de latitude e longitude do(s) ponto(s) selecionado(s) permite que o usuário se certifique dos parâmetros da seleção antes de executar uma busca na base de dados (“Informações do Mapa”). O formulário para seleção de dados do BDG é apresentado à direita e logo abaixo são apresentadas as informações resultantes de uma consulta realizada pelo usuário.

Para realizar uma busca (“Buscar Dados”) deve-se informar o período para qual deseja visualizar os dados de precipitação. A série histórica possui registros de precipitação da grade com resolução temporal de três horas. Assim, deve-se fornecer também a resolução temporal da saída desejada, que pode variar entre 03 horas até 01 dia.

Após a execução das buscas, o usuário pode gerar um arquivo de saída (“Guardar Busca”) contendo o resultado obtido, o que permite o emprego desses dados em outros programas, como por exemplo, um SIG ou um sistema estatístico.

O sistema não se limita aos pontos utilizados neste estudo de caso. Para adicionar pontos de outros locais de interesse, basta adicionar os dados de precipitação no BDG do sistema.

## 5. Conclusão

O uso de SIGs *web* por profissionais que tomam decisões no monitoramento ambiental é uma solução de acesso fácil, rápido e de baixo custo (SCHIMIGUEL *et al.*, 2014). Neste contexto o sistema desenvolvido facilita o manejo dos dados de precipitação fornecidos pelo TRMM, por meio de uma interface simples e intuitiva.

No processo de busca, buscou-se minimizar o trabalho do usuário, podendo este escolher um ponto em específico ou selecionar vários pontos do mapa com o *mouse*. Buscou ainda proporcionar ao usuário maior flexibilidade na ação de seleção dos pontos, permitindo também a entrada com o número identificador do(s) ponto(s).

O sistema desenvolvido neste trabalho é totalmente funcional, podendo ser hospedado em um servidor para utilização imediata, como suporte às pesquisas que necessitam dos dados de precipitação fornecidos pelo projeto TRMM.

## 6. Referência Bibliográfica

KUMMEROW, C. et al. The status of the Tropical Rainfall Measuring Mission (TRMM) after two years in orbit. *Journal of Applied Meteorology*, v. 39, p.1965-1982, 2000.

MACEDO, M. J. H. Aplicações De Redes Neurais Artificiais e Satélite TRMM na Modelagem Chuva-Vazão da Bacia Hidrográfica do Rio Paraguaçu/BA. [s.l.] Tese de Doutorado. Universidade Federal de Campina Grande, junho 2013.

NASA. Disponível em: <<http://trmm.gsfc.nasa.gov>>. Acesso em: 29 de julho de 2014.

NÓBREGA, R.S.; SOUZA, E.P.; GALVÍNCIO, J.D. Análise da estimativa de precipitação do TRMM em uma sub-bacia da Amazônia Ocidental. *Revista de Geografia*, v. 25, p. 6-20, 2008.

PERETTI, A. V.; NUMMER, A. V. Desastres naturais ocorridos no município de Erechim/RS entre os anos de 1986 a 1991. *Revista Geonorte*, v.1(Edição Especial), p.591–598, 2012.

SCHIMIGUEL, J; BARANAUSKAS, M. C.; MEDEIROS, C. B. Usabilidade de aplicações SIG *web* na perspectiva do usuário: um estudo de caso. Disponível em: <[http://www.ip.pbh.gov.br/ANO8\\_N1\\_PDF/ANO8N1\\_Juliano.pdf](http://www.ip.pbh.gov.br/ANO8_N1_PDF/ANO8N1_Juliano.pdf)>. Acesso em 27 de setembro de 2014.

## Aplicação para busca de similaridade em séries temporais

Luy Lucas Ribeiro Santana, Allan G. de Oliveira, Josiel Maimone de Figueiredo,  
Thiago Meirelles Ventura, Claudia Aparecida Martins

Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)

Cuiabá – MT – Brasil

luylucas10@gmail.com, {allan, josiel, thiago, claudia}@ic.ufmt.br

**Abstract.** *This work aims to present the development of a software to usage of a similarity search engine with application in micrometeorological data. The software has been developed in order to assist users, experts and researchers in data interpretation and decision making using this data.*

**Resumo.** *Este trabalho tem como objetivo apresentar o desenvolvimento de um software para utilização de técnicas de busca por similaridade com aplicação em dados micrometeorológicos. A aplicação vem sendo desenvolvida com o intuito de auxiliar, tanto usuários quanto especialistas e pesquisadores na interpretação dos dados e tomada de decisões.*

### 1. Introdução

Séries temporais têm sido utilizadas em diversas áreas do conhecimento, que vão desde gerenciamento ambiental até dados de bolsa de valores. O uso desse tipo de dado tem crescido cada vez mais, principalmente pelo aumento da capacidade de aquisição de dados, obviamente, esse aumento implica em maior dificuldade de utilização desses dados para extração de informações úteis e tomada de decisões.

Nesse contexto, técnicas computacionais são desenvolvidas com o objetivo de auxiliar o usuário desses dados em sua interpretação. Dentre as diversas técnicas tem-se a busca por similaridade, que objetiva basicamente, encontrar padrões parecidos com um determinado trecho de série temporal. Infelizmente, apesar da existência de diversas técnicas, não há tantas ferramentas consolidadas disponíveis para utilização, ficando na maioria das vezes a encargo do usuário utilizar as soluções existentes independentemente.

Assim, o objetivo desse trabalho foi de construir uma ferramenta para realização de busca por similaridade em séries temporais, mais especificamente em dados meteorológicos. A utilização da técnica de busca por similaridade nesse tipo de dado se justifica devido à grande atenção que é dada atualmente para o controle ambiental, nesse sentido, tomadas de decisões podem ser auxiliadas identificando-se padrões

conhecidos em uma série temporal, como por exemplo, a identificação de um possível evento climático extremo.

## 2. Metodologia Utilizada

Para implementação do sistema para busca por similaridade em séries temporais utilizou-se a técnica de Janela Deslizante e inicialmente como métrica de similaridade o *Dynamic Time Warping* (DTW) (FU et al., 2005). O DTW utiliza uma função de distância para calcular a similaridade entre duas séries. Nesse trabalho foi utilizada a Distância Euclidiana.

### 2.1. Janela Deslizante

Dada uma série temporal  $S = S_1, S_2, \dots, S_n$ , uma janela de tamanho  $W$  e a sobreposição  $O$ . A janela deslizante extrai de  $S$  subséries  $S_i, \dots, S_{i+w}$ , a próxima subsérie é extraída por  $S_{i+o}, \dots, S_{i+w+o}$ . Este processo se repete até que  $i+w$  seja igual ao tamanho total da série. Cada subsérie extraída é comparada por meio da função de similaridade e as que obtiverem a menor distância são ditas similares a subsequência comparada.

### 2.2. Cálculo da Similaridade

A similaridade entre séries temporais é calculada pela distância entre as duas séries. A definição, no entanto, depende do usuário, do domínio e da tarefa a ser executada. (KEOGH, 2002). Usualmente, devido a demanda de processamento computacional as buscas não são realizadas na série inteira e sim em partes dela utilizando-se da técnica de janela deslizante. Para realizar a busca por similaridade deve-se primeiro definir a função de distância a ser utilizada. As funções de distância segundo Skopal (2011) devem seguir 4 postulados:

**Simetria:**  $D(c_1, c_2) = D(c_2, c_1)$ : a distância entre  $c_1$  e  $c_2$  deve ser a mesma para o oposto;

**Similaridade:**  $D(c_1, c_1) = 0$ : a distância do ponto à ele mesmo deve ser nula;

**Positividade:**  $D(c_1, c_2) = 0$  se, e somente se,  $c_1 = c_2$ ;

**Desigualdade triangular:**  $D(c_1, c_2) \leq D(c_1, c_3) + D(c_2, c_3)$ ;

A função de distância mais utilizada segundo Keogh & Kasetty (2002) e Last et al., (2004) é a distância euclidiana, descrita a seguir. Outras podem ser citadas, como a Manhattan (Hauswirt et al., 2005) e Euclidiana Quadrática (KEOGH, 2002).

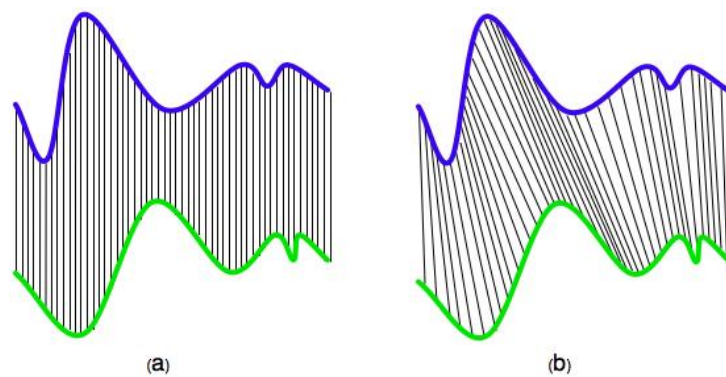
#### 2.2.1. Distância euclidiana.

Tendo duas séries  $A(a_1, a_2, \dots, a_n)$  e  $B(b_1, b_2, \dots, b_n)$ , a distância entre elas, utilizando a distância euclidiana, será dada pela equação:

$$D(A, B) = \sqrt{\sum_{i=1}^n (a_i - b_i)^2} \quad (1)$$

### 2.2.2. Dynamic Time Warping

A função de similaridade usada neste trabalho é o *Dynamic Time Warping* (DTW). Dado duas séries  $Q$  de tamanho  $n$  e  $C$  de tamanho  $m$ , onde os elementos são:  $Q = q_1 \dots q_n$  e  $C = c_1 \dots c_m$ , o DTW faz o alinhamento entre essas duas séries (figura 1) construindo uma matriz  $n$ -por- $m$  onde cada elemento  $M(i, j)$  da matriz corresponde a distância entre os pontos  $q_i$  e  $c_j$ . Uma observação: no caso desta aplicação só é possível utilizar séries de tamanhos iguais.



**Figura 1 – comparação da Distância Euclidiana com o DTW: a) a Distância euclidiana realiza a distância de ponto à ponto. b) o DTW realiza um alinhamento entre as séries**

Um caminho  $K$  é traçado nesta matriz (figura 2) para encontrar o alinhamento entre as séries. Esse possui algumas propriedades:

**Limites:**  $k_1 = (1, 1)$  e  $k_n = (n, n)$ : isso força para que o caminho comece e termine em cantos diagonalmente opostos.

**Continuidade:** Sendo  $k_n = (a, b)$  então  $k_{n-1} = (a, b)$  onde  $a - a \leq 1$  e  $b - b \leq 1$ , isso restringe os possíveis passos do caminhamento para células que sejam adjacentes (incluindo diagonalmente adjacentes).

**Monotonicidade:** Sendo  $k_n = (a, b)$  então  $k_{n-1} = (a, b)$  onde  $a - a \geq 0$  e  $b - b \geq 0$ , isso força os pontos de  $K$  a serem monotonicamente espaçados no tempo.

Porém qualquer caminho  $K$  pode satisfazer essas propriedades interessa é a qual o custo é mínimo para a função, aquele ao qual representará, mas a que a menor distância entre as séries:

$$DTW(A, B) = \min\{\sqrt{\sum_{a=1}^K k_a}\} \quad (2)$$

Este caminho pode ser encontrado utilizando programação dinâmica para buscar os próximos elementos do conjunto  $K$ . Assim, a distância cumulativa definida pela

programação dinâmica  $\gamma(i, j)$  e o mínimo das distâncias cumulativas das células adjacentes, podem ser expressadas como:

$$\gamma(i, j) = d(a_i, b_j) + \min\{\gamma(i-1, j-1), \gamma(i-1, j), \gamma(i, j-1)\} \quad (3)$$

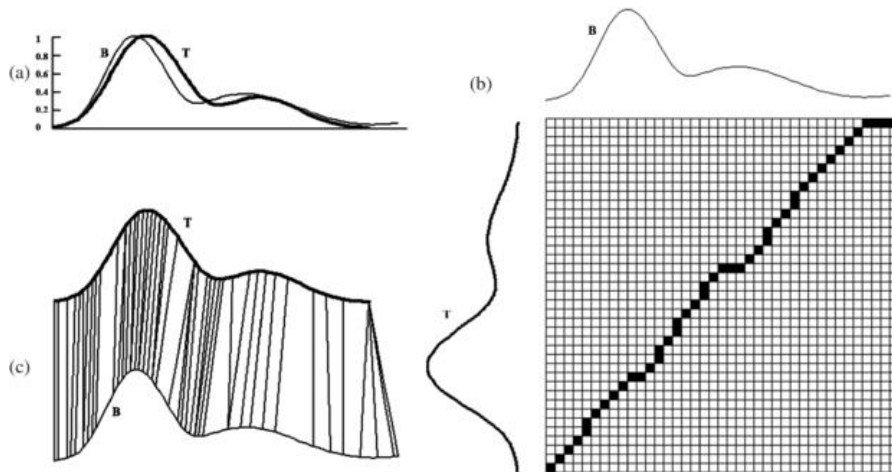


Figura 2 - a) duas séries sobrepostas nos eixos X e Y. b) matriz com as distâncias. Cada elemento  $(i, j)$  representa a distância entre os próprios elementos. O caminho traçado representa o menor custo ou também a menor distância entre as séries. c) alinhamento com o DTW entre das duas séries.

### 3. Descrição do software

O objetivo do desenvolvimento do software foi basicamente fornecer uma interface simples para a execução da tarefa de busca por similaridade para dados meteorológicos. A implementação foi feita utilizando a linguagem de programação Java, utilizou-se a API JavaFX para desenvolver a interface gráfica. A escolha do JavaFX foi feita por que a API permite que uma única implementação seja feita para aplicação *desktop* e *web*.

A base do sistema de busca parte da definição das classes que operam o sistema aparece na figura 3. Foi definida uma classe interface chamada *SimilarityFunction* com o objetivo de que novos métodos possam ser adicionados, bastando que o mesmo implemente o método *getSimilarity* declarado na interface. A estrutura de classes foi pensada de forma com que seja possível adicionar novas funções de similaridade e novas funções de distância. Nesse trabalho, como ilustrado na figura 3, foi implementado como medida de similaridade o DTW que nesse caso utiliza como função de distância a distância euclidiana.

Além da estrutura de classes para a execução dos métodos de similaridade, o objetivo foi de implementar uma interface gráfica que facilite o uso das técnicas por pesquisadores da área. A Figura 4 mostra uma tela do sistema com uma série de dados e ao lado direito da tela uma tabela com alguns resultados de uma busca realizada.

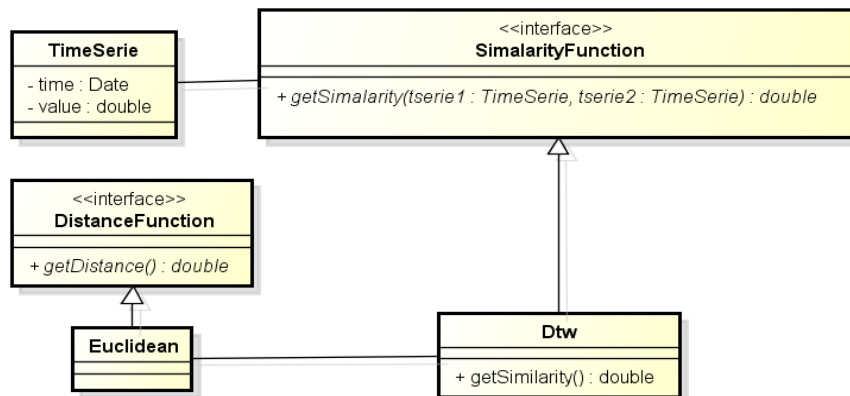


Figura 3 Diagrama de Classes base do sistema.

Nessa tabela as informações são ordenadas em ordem crescente de similaridade. É possível ver o início e final do período e a similaridade encontrada e ainda através de *check-box* na tabela selecionar quais os períodos deseja-se plotar no gráfico para comparar com a série original.

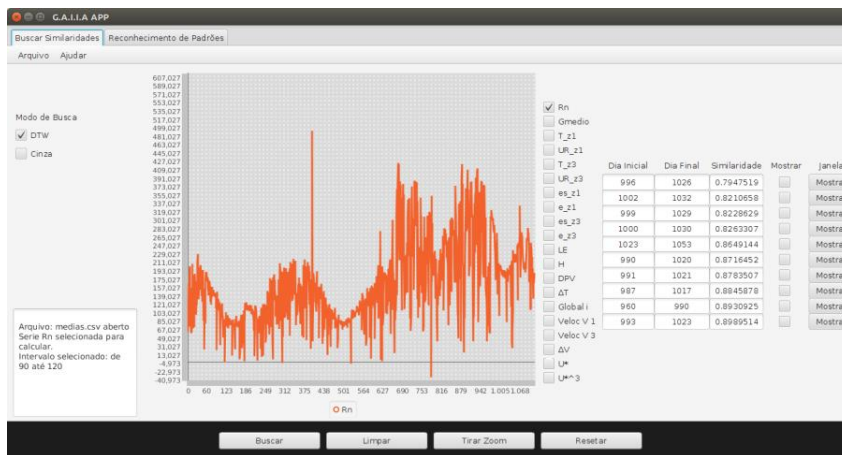


Figura 4 Interface da aplicação.

Na Figura 5 é mostrada uma aproximação do centro da tela que exibe os dados e ainda uma lista lateral que contém outros dados. Essa funcionalidade é útil quando o arquivo de dados utilizado pelo usuário contém mais de um tipo de dado e deseja visualizar mais de um no mesmo gráfico e ainda fazer a busca por similaridade em séries de dados diferentes.

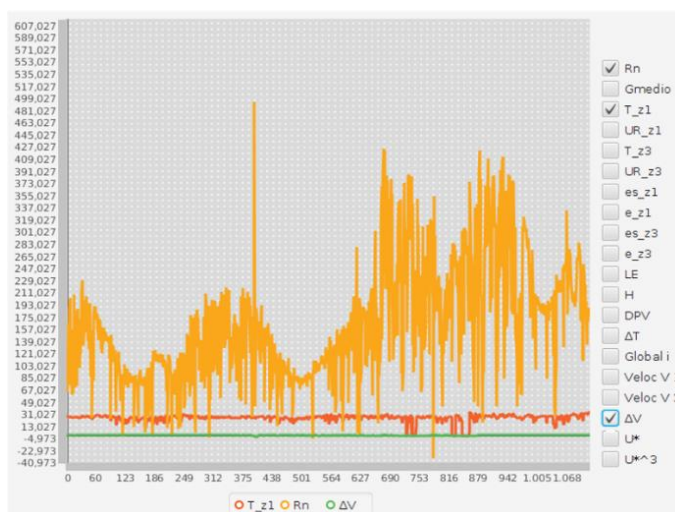


Figura 5 - Visualização de mais de uma série no mesmo gráfico.

#### 4. Conclusão

Este trabalho apresentou o desenvolvimento de um software para a aplicação e desenvolvimento de técnicas de busca por similaridade com aplicação em dados micrometeorológicos. O desenvolvimento de uma aplicação com interface simples para o usuário especialista nesse tipo de dado pode ampliar a utilização desse tipo de técnica e pode contribuir com as pesquisas que utilizam dados micrometeorológicos.

#### 5. Referências

FU, A. W. et al. Scaling and Time Warping in Time Series Querying. **Very Large Data Base Journal**. v.17, p.899-921, 2005.

Hauswirt, M., Diwan, A., Sweeney, P. F, Mozer, M. C. Automating vertical profiling. In Proceedings of the 20th Annual Conference on Object Oriented Programing, Systems, Languages, and Applications, pág. 281-296, New York, USA. ACM Press, 2005.

Keogh, E. Ratanamahatana C. A. **Exact indexing of dynamic time warping**. In Proceedings of the 28th international conference on Very Large Data Bases. Pág. 406-417. 2002.

Keogh, E., Kasetty, S. Constructive induction for classifying time series. In Proceedings of the 15th European Conference on Machine Learning, volume 3201 of Lecture Notes in Computer Science, pág. 192-204, Pisa, Italy. Springer. , 2002.

Skopal, T., Bustos, B. 2011. On nonmetric similarity search problems in complex domains. *ACM Comput. Surv.* 43, 4. 2011.

# Análise da capacidade de comunicação dos sensores Rádiuino

Franco Magri Lupo<sup>1</sup>, Ivairton Monteiro Santos<sup>1</sup>

<sup>1</sup>Instituto de Ciências Exatas e da Terra – Campus Universitário do Araguaia –  
Universidade Federal de Mato Grosso (UFMT)  
Barra do Garças – MT – Brasil

franco.lupo@gmail.com, ivairton@ufmt.br

**Abstract.** *This work perform an evaluation of the communication capacity of Rádiuino devices. This devices is used to compose a wireless sensor networks. The tests was used the devices in your default configuration, without routing data algorithms, or collision avoidance. The aim was verify the max length of communication between sensor nodes, the receiving capacity of the mobile gateway node and the interference of data packet collision in communication. The results showed that the packet data collision is the main aspect to be considered in communication between Rádiuino devices.*

**Resumo.** *Este trabalho executa uma avaliação da comunicação dos dispositivos Rádiuino, utilizado para compor redes de sensores sem fio. A análise foi feita com os dispositivos em sua configuração original, sem a utilização de algoritmos de roteamento de dados ou de prevenção de colisão de pacotes. Os testes objetivaram avaliaram o alcance máximo de comunicação entre dois dispositivos, a capacidade de recepção de pacotes de dados com o nó gateway em movimento e o impacto na qualidade da comunicação dos nós sensores causado pela colisão de pacotes de dados. Os testes demonstraram que a colisão dos pacotes é o principal fator atenuador da qualidade de comunicação entre dispositivos Rádiuino.*

## 1. Introdução

As redes de sensores sem fio (RSSF) vem sendo cada vez mais utilizadas para o monitoramento e automação de sistemas, com aplicação em diversas áreas (Akyildiz *et al.*, 2002). Com os avanços tecnológicos, há no mercado inúmeras opções de dispositivos para o desenvolvimento de projetos de RSSF (Hill *et al.*, 2004). Uma dessas opções é o Rádiuino (Branquinho, 2011), baseado na arquitetura Arduino (Banzi, 2009), que tem como objetivo principal a criação de projetos e protótipos de RSSF.

Para o desenvolvimento adequado de um projeto e implantação de uma RSSF é preciso conhecer as capacidades e limitações dos dispositivos a serem utilizados. Este trabalho irá avaliar a qualidade da comunicação dos dispositivos Rádiuino em conjunto com o rádio transmissor BE900 (disponível para utilização pelo grupo de pesquisa) (Radioit, 2012), considerando diferentes contextos organizacionais, empregando os sensores Rádiuino em sua configuração original.

Os testes buscaram avaliar o alcance de comunicação, a interferência de obstáculos na comunicação, a influência na comunicação quando há o emprego de um

*gateway* móvel e o quanto a colisão de pacotes inviabiliza na comunicação dos nós sensores da rede, formada pelos componentes descritos.

## 2. Metodologia dos testes

Os experimentos buscaram avaliar a eficiência da comunicação entre sensores Radiuino homogêneos, todos em suas configurações originais. A Figura 1 ilustra os componentes empregados nos testes: suporte para baterias (à esquerda na figura), módulo Radiuino modelo DK102 v12 (à direita na figura), rádio transmissor modelo BE900 (na parte superior da figura) e módulo UartBee (no centro da figura) para gravação nos microcontroladores e comunicação com o computador (por meio de um nó *gateway*, que corresponde ao nó receptor dos pacotes de dados enviados pelos nós da rede de sensores).



Figura 1. Foto de um *kit* Radiuino modelo DK102.

Foram executados quatro tipos de testes: 1) alcance de comunicação; 2) interferência de obstáculos; 3) capacidade de comunicação dos nós sensores com um *gateway* móvel; e 4) perda de pacotes por colisão.

### 2.1. Alcance de comunicação

O primeiro teste consiste em avaliar o alcance de comunicação entre os nós sensores Radiuino. É importante observar que os sensores utilizados não possuem antena de comunicação, há apenas um fio soldado na placa.

O teste consiste no envio de pacotes de dados por um nó sensor fixo para o nó *gateway*, posicionado em um local aberto, sem a presença de obstáculos. Gradativamente o *gateway* foi se distanciando do nó transmissor até que a taxa de recebimento de pacotes ficou abaixo dos 5%. Foram utilizadas oito configurações diferentes para a potência do sinal, variando da mais fraca para a mais forte: -30 dbm, -20 dbm, -15 dbm, -10 dbm, 0 dbm, 5 dbm, 7 dbm e 10 dbm. O alcance de comunicação foi aferido por meio de GPS.

### 2.2. Interferência na comunicação por obstáculos

Esse teste buscou avaliar o quanto obstáculos podem interferir no alcance de comunicação dos nós sensores em relação à comunicação em campo aberto.

A estratégia empregada neste testes foi semelhante à utilizada no teste (anterior) de alcance de comunicação. No entanto, apenas a potência do sinal máxima (10 dbm) foi utilizada. O obstáculo utilizado no teste caracteriza-se exclusivamente por vegetação, não densa, composta por arbustos e árvores de pequeno porte. No processo de afastamento do *gateway*, além do incremento da distância, a densidade do obstáculo também aumentou.

### 2.3. Recepção de pacotes de dados com o gateway em movimento

Nesse ensaio a mobilidade do gateway foi posta à prova, de modo a verificar se o seu movimento interfere na recepção dos pacotes de dados enviados pelos nós sensores.

Com o nó gateway instalado em um veículo e conectado a um notebook, foram executadas passagens a 20, 30, 40, 50 e 60 km/h. Os nós sensores (fixos) foram dispostos ao longo de uma via em três configurações de quantidade distintas, com 10, 5 e 3 nós. Na configuração com 10 nós, os sensores foram dispostos em duas linhas com 5 unidades, onde a distância entre os nós ao longo da linha era de 15 m e entre as linhas era de 9 m. O gateway movimentou-se ao longo do corredor formado. Na configuração com 5 unidades, permaneceu apenas uma das linhas descrita anteriormente. Na configuração com 3 nós sensores, permaneceram apenas os sensores das extremidades e o central, em relação à configuração com 5 nós. Portanto, a distância entre os sensores foi de 30 m. A Figura 2 apresenta fotos do gateway fixado no automóvel e de um nó sensor disposto na via (sobre a sarjeta).



Figura 2. Dispositivo fixado no automóvel, para simulação de um gateway móvel (à esquerda). Nó sensor posicionado na via (à direita).

Outros dois parâmetros foram trabalhados nesse teste. A potência do sinal de rádio comunicação, tendo sido utilizado apenas os valores mínimo (-30 dbm) e máximo (10 dbm); e o intervalo entre o envio de cada pacote de dado pelos nós sensores. Este intervalo foi de 0,2 s ou 0,5 s. Vale destacar que não houve nenhum tipo de sincronização entre os nós sensores.

### 2.4. Perda de pacotes por colisão

Para se ter melhor compreensão da influência da colisão de pacotes de dados na comunicação entre os sensores, esse teste utilizou 5 nós sensores configurados em potência de comunicação de rádio máxima (10 dbm). Foram utilizadas duas configurações de frequência de envio dos dados, uma frequência com envio de pacotes a cada 0,2 s e outra com frequência de um pacote a cada 0,5 s.

Nesse teste foram utilizados 5 nós sensores, distantes 5 m entre si e todos a 15 m do gateway, em campo aberto, sem a interferência de obstáculos. O teste consistiu em monitorar o recebimento dos pacotes de dados de cada nó sensor pelo gateway pelo período de 1 minuto, registrando também o número de pacotes enviados por cada nó sensor. No primeiro estágio do teste apenas 1 nó sensor enviou dados. Na sequência, um nó sensor foi ligado e novamente executado o teste, até que no último estágio, os 5 nós sensores estavam enviando pacotes de dados durante o intervalo de tempo do teste. É importante destacar que não houve nenhum tipo de sincronização entre os nós sensores.

### 3. Resultados

Todos os testes foram executados por mais de uma vez, de modo que os resultados apresentados correspondem a uma média dos valores obtidos.

#### 3.1. Alcance de comunicação

A Figura 3 apresenta o gráfico com os resultados obtidos no teste de alcance de comunicação, sem obstáculos. Nota-se que utilizando a potência mínima no rádio transmissor o alcance máximo obtido foi de 21,5 m, enquanto que em potência máxima o alcance chegou a 205 m.

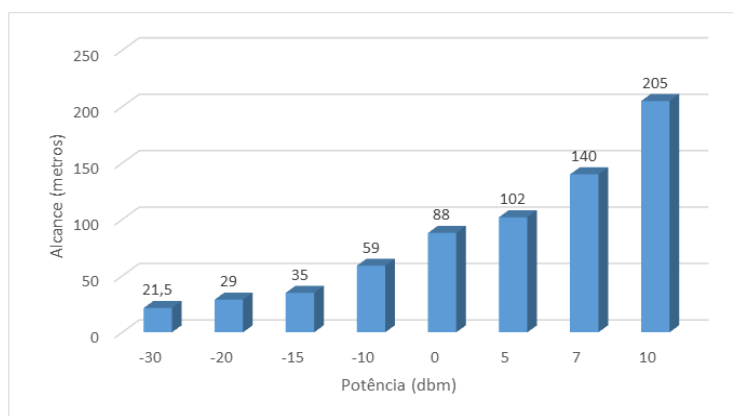


Figura3. Gráfico com os resultados do alcance de comunicação dos nós sensores, variando a potência do sinal de rádio.

#### 3.2. Interferência na comunicação por obstáculos

Conforme esperado, os resultados desse teste demonstraram que a existência de obstáculos pode interferir de maneira substancial na qualidade de comunicação dos sensores Radiuno. Enquanto que o alcance de comunicação entre dois nós sensores, configurados com potência máxima no sinal de rádio chegou a 205 m, no contexto com obstáculos (vegetação) essa mesma configuração obteve 80 m. O resultado obtido com a existência de obstáculos corresponde a 39% da capacidade de alcance em campo aberto.

Essa interferência pode ser explicada pelo fato dos sensores trabalharem em frequência de 2,4 Ghz, sabendo-se que esta frequência é sensível à água e que a vegetação possui considerável índice de água, tal contexto justifica a queda na capacidade de alcance de comunicação. Além ainda, da possível ocorrência de refração.

#### 3.3. Recepção de pacotes de dados com o gateway em movimento

Nesse teste foram executadas 5 passagens do gateway pela rede de sensores para cada contexto (combinação das configurações descritas na Seção 2.3).

O teste registrou o número de pacotes de dados enviados por cada nó sensor a partir do instante que o gateway entrou na sua área de alcance de comunicação, registrando também o número e a origem de cada pacote de dado recebido pelo gateway. Dessa forma, foi calculada a média da razão entre o número de pacotes recebidos pelo gateway, pelo número de pacotes enviados, com objetivo de avaliar a porcentagem de pacotes de dados recebidos em cada contexto.

A Figura 4 apresenta o gráfico com os resultados obtidos para os testes feitos com 10 nós sensores, enquanto que a Figura 5 o gráfico com os resultados para a rede com 5 nós sensores e a Figura 6 o gráfico com os resultados para a rede com 3 sensores.

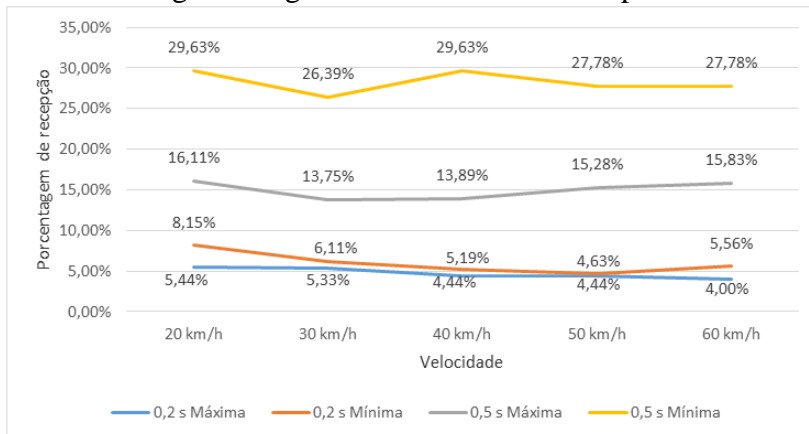


Figura4. Porcentagem de pacotes recebidos pelo gateway móvel da rede com 10 nós.

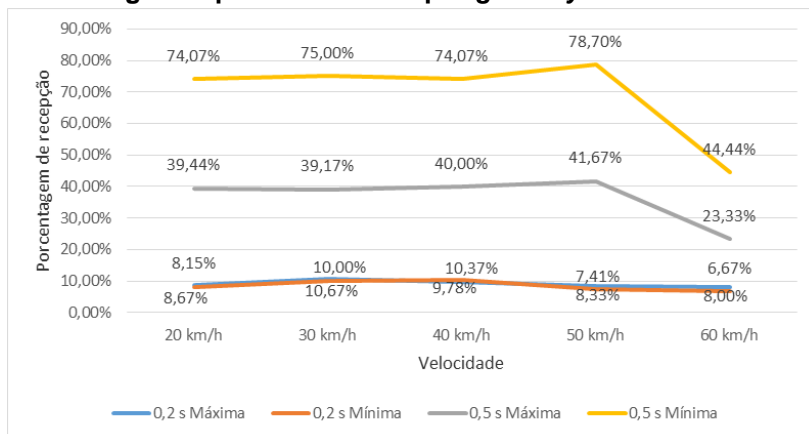


Figura5. Porcentagem de pacotes recebidos pelo gateway móvel da rede com 5 nós.

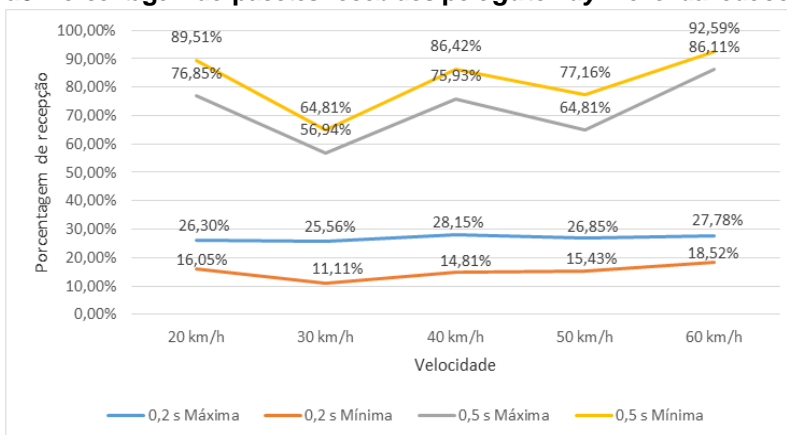


Figura6. Porcentagem de pacotes recebidos pelo gateway móvel da rede com 3 nós.

Os gráficos, dos diferentes testes, mostraram que não há uma redução na taxa de recebimento de pacotes de dados pelo gateway móvel em relação às velocidades testadas.

No entanto, observa-se indícios que a ocorrência de colisão de pacotes interferiu de maneira razoável nos resultados obtidos. Por exemplo, as melhores taxas de entrega

sempre foram para configurações dos sensores com menor frequência de envio de pacotes e a configuração da rede com menor número de nós obteve melhores taxas do que os outros contextos.

### 3.4. Perda de pacotes por colisão

Esse teste evidenciou que a colisão de pacotes é fator determinante na qualidade de comunicação entre os nós sensores. A Figura 7 apresenta o gráfico com os resultados obtidos no teste. Observa-se que com o funcionamento de apenas 1 nó sensor, todos os pacotes enviados foram recebidos, enquanto que no funcionamento de 5 nós sensores em paralelo, essa taxa chegou ao índice de apenas 10,25%.

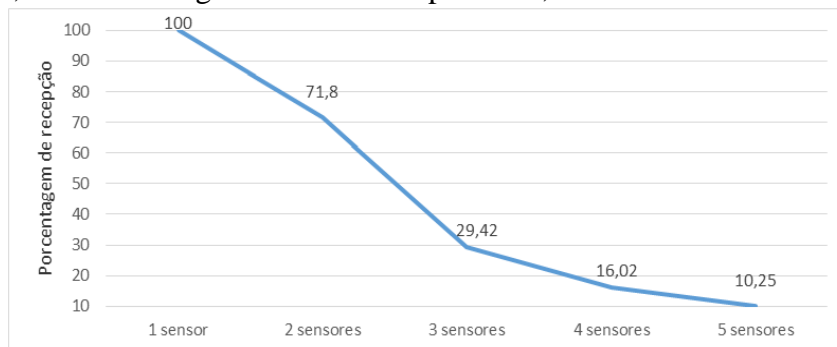


Figura 7. Porcentagem de pacotes recebidos pelo gateway para rede com diferentes números de sensores.

## 4. Conclusão

Este trabalho buscou avaliar as características de comunicação de uma rede de sensores sem fio formada por componentes Rádiumino, em sua configuração original. Os testes demonstraram que o alcance de comunicação máximo em campo aberto e sem obstáculos pode chegar até a ~200 m. Se o contexto da aplicação houver obstáculos formados por vegetação esse alcance cai para ~80 m.

Verificou-se também que a mobilidade do *gateway*, com velocidade de até 60 km/h não interfere na capacidade de recepção dos pacotes de dados. No entanto, a colisão de pacotes foi determinante no funcionamento da rede de sensores. Dessa forma, é necessário empregar algoritmos de prevenção de colisão de pacotes.

## References

- Akyildiz, I. F., Su, W., Sankarasubramaniam, Y., and Cayirci, E. (2002). “Wireless sensor networks: A survey”, In: *Computer Networks*, v. 38, pp. 393–422.
- Banzi, M. (2009). *Getting started with Arduino*, O'Reilly, 1ª edition.
- Branquinho, O. (2011). *Plataforma Rádiumino para estudos em redes de sensores sem fio*. Acesso online <<http://www.radiumino.cc>>.
- Hill, J., Horton, M., Kling, R. E Krishnamurthy, L. (2004). “The Platforms Enabling Wireless Sensor Networks”, In: *Commun. ACM*, v. 47, n. 6, pp 41–46.
- Radioit (2012). *BE 900 Datasheet*. Acesso online <<http://www.radioit.com.br/sites/default/files/downloads/BE900 Datasheet.pdf>>.

## A Computação em Nuvem e os Desafios de Infraestrutura para e-Science

David Beserra<sup>1</sup>, Felipe Fernandes<sup>2</sup>, Francielly Falcão<sup>1</sup>, Wagner Lima<sup>1</sup>, Rubens Karman<sup>3</sup>, Alexandre Borba<sup>1</sup>, Jean Araujo<sup>1</sup>, Erica Sousa<sup>1</sup>

<sup>1</sup>Unidade Acadêmica de Garanhuns – Universidade Federal Rural de Pernambuco  
Garanhuns – PE – Brasil

<sup>2</sup>Pós-graduação em Sistemas e Computação – Instituto Militar de Engenharia  
Rio de Janeiro – RJ – Brasil

<sup>3</sup>Escola de Computação Inteligente – Universidade de Pernambuco  
Recife – PE – Brasil

**Abstract.** *This paper presents the infrastructure challenges inherent to the full development of e-Science applications. This paper outlines how cloud computing can collaborate to improve the infrastructure provisioning to e-Science application support. We also show the main open source tools that can be used in the cloud implementation in order to improve the high performance computing in e-Science.*

**Resumo.** *Este artigo apresenta os desafios de infraestrutura inerentes ao pleno desenvolvimento de aplicações de e-Science. Este artigo descreve como a computação em nuvem pode colaborar para melhorar o fornecimento de infraestrutura para suporte de aplicações de e-Science. São mostradas também ferramentas de código aberto que podem ser usadas na implementação de nuvens que proporcionam a computação de alto desempenho em e-Science.*

### 1. Introdução

Diversas áreas da ciência demandam por sistemas de Computação de Alto Desempenho (CAD). Inicialmente, supercomputadores foram ofertados para atender esta demanda, entretanto seus custos de aquisição os tornam inviáveis para muitas instituições de ensino. Como alternativa de baixo custo surgiu o *Cluster Beowulf*. Entretanto, mesmo tendo sido um avanço, *clusters* isolados não mais atendem as necessidades da computação científica, reflexo da tendência de integração entre instituições para cooperação em projetos de grande porte. Logo, passou-se a requerer a integração entre os sistemas de CAD isolados sobre uma única matriz distribuída geograficamente como forma de aumentar os recursos computacionais para resolver em tempo hábil problemas cada vez mais complexos.

Uma alternativa inicial encontrada pelas instituições foi a computação em grade (ou *Grid Computing*), que pode ser comparada metaforicamente com a rede elétrica [Cirne 2003], onde os equipamentos apenas são conectados e recebem a energia necessária. De maneira similar, pode-se visualizar um *grid* como sendo uma rede sobre

a qual os usuários se conectam para obter serviços [Cirne 2003]. Particularmente, no caso da computação científica, os serviços ofertados e obtidos são de processamento de alto desempenho, armazenamento e compartilhamento de bancos de dados (BD).

Como evolução ao *grid* surgiu a computação em nuvem (*Cloud Computing*), que sugere a si próprio como um sistema de computação distribuída de larga escala, oferecendo um conjunto de recursos virtualizados, dinamicamente escalonáveis e facilmente gerenciáveis [Liu et al. 2011]. Em uma nuvem, diversos serviços podem ser disponibilizados simultaneamente sobre o mesmo *hardware*, independente de suas características computacionais, tal como *clusters* ou *grids*. Este paradigma, quando ofertado sobre a forma de IaaS, melhora o apoio a ciência em larga escala (*e-Science*) [Mattoso et al. 2008].

Neste trabalho é descrito como a computação em nuvem pode ser utilizada para prover infraestrutura computacional para CAD aplicada à *e-Science*. A seguir serão apresentados, na Seção 2, os conceitos principais referentes a *e-Science* e os seus desafios de infraestrutura. Na seção seguinte, como a Computação em Nuvem pode auxiliar na superação dos desafios atuais de infraestrutura para *e-Science*. Na Seção 4, algumas soluções *open source* para a implementação de nuvens IaaS. Na Seção 5, as conclusões e as ideias para trabalhos futuros.

## 2. Desafios de Infraestrutura em e-Science

*“Fazer ciência hoje implica, dentre outros aspectos, ubiquidade e distribuição, visando ao desenvolvimento e execução de soluções com alto desempenho, baseadas em reutilização, gerência de dados e experimentos. Daí surge o termo e-Science, que caracteriza o apoio a ciência em larga escala utilizando infraestrutura computacional”* [Mattoso et al. 2008].

O conceito de *e-Science* deve permitir o acesso aos recursos computacionais distribuídos, ao armazenamento e ao compartilhamento de dados e resultados, para viabilizar o desenvolvimento de pesquisas colaborativas por equipes internacionais e interinstitucionais, separadas por grandes distâncias [Mello et al. 2010]. Sistemas de gerência de *workflows* científicos foram desenvolvidos para apoiar a gerência da composição de processos e dados num fluxo que possa encadear as diversas análises de dados, contudo, o apoio computacional ao experimento científico em larga escala ainda não é satisfatório [Gil et al. 2007].

A otimização no uso dos recursos em centros de CAD e sua integração com a infraestrutura de *e-Science* apresentam desafios, como a necessidade por sistemas de *software* mais eficientes. É necessário desenvolver programas com melhor escalabilidade e eficiência temporal, que aproveitem melhor os sistemas multiprocessados [Riedel et al. 2007]. Todavia, o maior desafio é a oferta de infraestrutura bruta de processamento escalonável dinamicamente e a integração entre as diversas estruturas existentes, de forma que possam ser utilizadas como um recurso unificado, o que pode ser conseguido mediante o modelo da Computação em Nuvem sob a ótica da oferta de IaaS.

Foi exposto por [Keahey et al. 2005] a ideia de *Virtual Workspaces*, onde é previsto o provisionamento automático de recursos como maneira de atender as

necessidades dos usuários, sendo este conceito de total adequação a *e-Science*. Este conceito também entra em sintonia direta com as definições de nuvem, que mantém alocação dinâmica de recursos e elasticidade como intrínsecos a si.

Entretanto, é desafiador prover um ambiente de *hardware* e *software* que ofereça a seus usuários abstração sobre os recursos computacionais necessários à execução de suas aplicações, particularmente simulações. Similarmente, prover uma infraestrutura de gerência dos elementos computacionais utilizados durante estas simulações também apresenta-se como um processo complexo. Além disso, o avanço científico exige investimento crescente em equipamentos e infraestrutura, de forma a suprir as necessidades de tal avanço [Mello et al. 2010]. Se ocorrer indisponibilidade infraestrutural, as pesquisas científicas efetuadas por um usuário podem não fluir em um ritmo adequado, comprometendo sua competitividade frente ao progresso da ciência no cenário mundial [Mattoso et al. 2008]. Sob esta ótica, as tecnologias de virtualização e o paradigma de Computação em Nuvem passam a ter atuação decisiva na construção de infraestruturas adequadas para a superação dos desafios de *e-Science*.

### 3. Computação em Nuvem e e-Science

Existem entidades que estabelecem relacionamentos entre si em um ambiente de nuvem. [Gonçalves et al. 2011] identifica três entidades, as quais denominam-se agentes, que operam sobre um cenário típico de nuvem computacional. O primeiro destes é responsável pelo fornecimento dos serviços e é identificado como Provedor. Os outros dois agentes são denotados Cliente e Desenvolvedor. Neste cenário um Desenvolvedor pode ser um cliente do Provedor, assim como o próprio Cliente, que consome os serviços. O Desenvolvedor também consome serviços dos Provedores e os disponibiliza aos Clientes por meio da nuvem. Um esquema simplificado dos relacionamentos em uma nuvem pode ser visto na Figura 1.

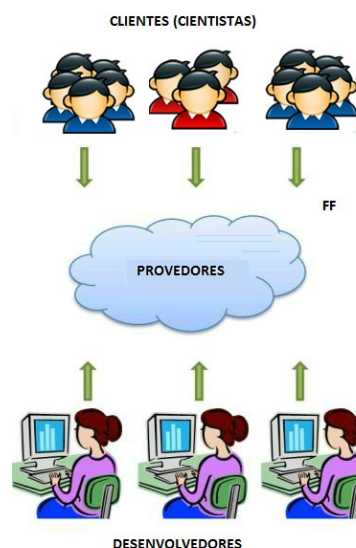


Figura 1. Relacionamentos em ambiente de nuvem

Analogamente em *e-Science*, em um cenário típico os Clientes seriam os cientistas que necessitam de infraestrutura computacional para execução de

experimentos. Os Desenvolvedores seriam todos os profissionais que realizam o suporte aos referidos experimentos, podendo os cientistas em questão também estarem inseridos nesta categoria, sendo eles próprios Clientes e Desenvolvedores. Neste cenário, o papel do Provedor caberia ao conjunto de instituições as quais os pesquisadores de um dado projeto estão inseridos. Logo, as instituições passam a ser provedores de IaaS para todo este conjunto de pesquisadores.

#### **4. Ferramentas para Implementação de Nuvem do tipo IaaS**

São expostas aqui 4 soluções de código aberto que são utilizadas em projetos para atender as demandas da comunidade de *e-Science*.

##### **4.1. Eucalyptus**

O Eucalyptus permite a criação de nuvens privadas e híbridas focando em IaaS, acessíveis via uma API compatível com os serviços da Amazon, como o EC2 e o S3. Existem quatro componentes em sua arquitetura. No nível mais interno está o *Node Controller* (NC), que é executado nos servidores hospedeiros de máquinas virtuais (VMs). Acima do nível do NC atua o *Cluster Controller* (CC), que atua agrupando os NCs em *clusters* e coletando informações de cada NC dentro do seu domínio de atuação. No nível mais externo temos o *Cloud Controller* (CLC), que toma decisões sobre onde alocar VMs. O Eucalyptus tem sido utilizado como gerenciador de infraestrutura em muitas aplicações para *e-Science*. Um exemplo é dado em [Kijispongse e Vannarat 2010], que implementou uma arquitetura de provisionamento automático de recursos para *Clusters Beowulf* por meio de recursos provenientes de uma nuvem Eucalyptus.

##### **4.2. OpenNebula**

O OpenNebula é utilizado como uma plataforma aberta para suprir os desafios de pesquisa encontrados na academia e na indústria. Sua arquitetura é semelhante à estrutura de um *cluster*, possuindo um servidor o papel de *frontend* e alocando nos demais servidores todas as VMs que forem necessárias. Para a integração com redes externas, provê mecanismos para o suporte de requisições provenientes de nuvens EC 2 e de nuvens implementadas com Eucalyptus [Sotomayor et al. 2009], adequando-se a necessidade de integração entre as diversificadas infraestruturas da academia. Em [Fernandes, Schulze, e Mury 2010] é apresentada uma proposta de um ambiente virtual de trabalho para uso em unidades de pesquisa, usando o OpenNebula como ferramenta de nuvem, que corrobora com o conceito de *Virtual Workspaces* presente em *e-Science*.

##### **4.3. Nimbus**

O Nimbus é focado em prover IaaS para atender especificamente a comunidade científica. Ele se baseia em três pontos principais: permitir a implementação de nuvens públicas ou privadas como IaaS; habilitar os usuários a usar nuvens IaaS e habilitar os desenvolvedores de aplicações a experimentar, customizar e estender nuvens IaaS. A extensão da nuvem se dá mediante o uso de interfaces com a Amazon EC2. Existem iniciativas que empregam o Nimbus no fornecimento de IaaS para *e-Science*, como o projeto Science Clouds [Science Clouds 2013], que provê um grupo de nuvens de pequeno porte para *e-Science* desde 2008.

#### 4.4. OpenStack

O OpenStack é uma ferramenta para implementação de nuvens computacionais com três serviços principais: *OpenStack Compute Infrastructure* (Nova), *OpenStack Storage Infrastructure* (Swift) e *OpenStack Imaging Service* (Glance) [Kureshi et al. 2013]. É utilizado em projetos estruturantes de *e-Science*, como o *QGG-cloud*, da Universidade de Huddersfield [Kureshi et al. 2013]. A *QGG-cloud* é uma nuvem privada para pesquisa em CAD. Está integrada a um *grid*. Tem implementada um mecanismo de alocação de recursos para *clusters* virtuais baseada em uma matriz de decisão, com propósito de maximizar a eficiência de utilização dos recursos da nuvem.

#### 5. Conclusões e Trabalhos futuros

O Quadro 1 apresenta um resumo das principais vantagens das ferramentas *Open source* apresentadas na Seção IV. O Eucalyptus, por possuir hierarquia estrutural similar à encontrada nos *Clusters Beowulf* é mais adequado para alocação de recursos destinados a CAD. O Open Nebula, por sua vez, apresenta vantagem quanto à integração de estruturas diversas, além de proporcionar *Virtual Workspaces*. O Nimbus tem como sua maior vantagem o fato de ser um projeto focado exclusivamente na comunidade científica, o que em perspectiva futura pode significar uma maior adaptação às necessidades da comunidade; é o único que conta com iniciativa especificamente direcionada à *e-Science* e seus requisitos de estruturas em nuvem. O OpenStack tem a facilidade de utilização como característica de grade vantagem.

**Quadro I. Características de ferramentas open source para nuvens IaaS aplicadas em e-Science**

| Ferramenta  | Vantagens   | Exemplos de utilização  |
|-------------|---|---|
| Eucalyptus  | Estrutura “ <i>cluster like</i> ”<br>Fácil gerenciamento                              | Provisionamento autônomo de recursos para <i>Clusters Beowulf</i>   |
| Open Nebula | Integração com Eucalyptus<br>Suporte a <i>Virtual Workspaces</i> ,                    | Integração de infraestruturas científicas distintas<br>Criação de múltiplos <i>Virtual Workspaces</i> distintos |
| Nimbus      | Foco na comunidade científica<br>Pequenas nuvens para <i>e-Science</i> já disponíveis | Projeto Science Clouds  |
| OpenStack   | Instalação simplificada   | <i>QGG-cloud</i>  |

Em contrapartida as vantagens, o custo-benefício de *clusters* para CAD alocados em uma nuvem por um período de tempo longo pode não ser atrativo [Napper e Bientinesi 2009]. Dado que a vantagem para CAD é a criação de sistemas sob demanda e não sistemas permanentes de uso geral, é importante verificar eventuais custos e riscos de integração entre nuvens distintas e entre nuvens e outras estruturas.

Como trabalhos futuros pretende-se investigar questões relativas ao desempenho das ferramentas apresentadas, para implantação de nuvens IaaS. Pretende-se também identificar o custo/benefício dessas ferramentas dado um determinado ambiente computacional pré-definido, apresentando as dificuldades, os benefícios e as justificativas da melhor solução para o ambiente em questão. Dessa forma, espera-se

que identificar pontos chaves na implementação de um ambiente de nuvem científica em um estudo de caso com validações e resultados comparativos ao ambiente anterior.

## Referências

- W. Cirne. (2003) "Grids computacionais: Arquiteturas, tecnologias e aplicações." III ERAD-Escola Regional de Alto Desempenho, Santa Maria, RS.
- F. Liu, J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger, e D. Leaf. (2011) "NIST Cloud Computing Reference Architecture." NIST Special Publication 500:292.
- T. C. Mello, B. Schulze, R. C. G. Pinto, A. R. Mury, (2010) "Uma análise de recursos virtualizados em ambiente de HPC", Anais do VIII Workshop em Clouds, Grids e Aplicações, pp. 17-30.
- M. Mattoso, C. Werner, G. H. Travassos, V. Braganholo, L. Murta. (2008) "Gerenciando Experimentos Científicos em Larga Escala", In: *Anais do XXVIII Congresso da SBC*, pp. 121-135.
- Y. Gil, E. Deelman, M. H. Ellisman, T. Fahringer, G. Fox, D. Gannon, C. A. Goble, M. Livny, L. Moreau, J. Myers. (2007) "Examining the Challenges of Scientific Workflows." *IEEE Computer*, v. 40(12), pp. 24-32.
- M. Riedel, T. Eickermann, S. Habbinga, W. Frings, P. Gibbon, D. Mallmann, F. Wolf, A. Streit, T. Lippert, W. Schiffmann, A. Ernst, R. Spurzem, e W. E. Nagel. (2007) "Computational steering and online visualization of scientific applications on large-scale hpc systems within e-science infrastructures." *eScience*, pp 483-490. IEEE Computer Society.
- K. Keahey, I. Foster, T. Freeman, and X. Zhang. (2005) "Virtual workspaces: Achieving quality of service and quality of life in the grid". *Sci. Program*, 13(4), pp. 265-275.
- G. E. Gonçalves, P. T. Endo, T. D. Cordeiro, A. V. A. Palhares, D. Sadok, J. Kelner, B. Melande, e J. Mångs. (2011) "Resource Allocation in Clouds: Concepts, Tools and Research Challenges". Anais do XXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos.
- Kureshi, I., Pulley, C., Brennan, J., Holmes, V., Bonner, S. e James, Y. (2013) *Advancing Research Infrastructure Using OpenStack*. International Journal of Advanced Computer Science and Applications, 3 (4). pp. 64-70.
- E. Kijispongse e S. Vannarat, S. (2010) "Autonomic resource provisioning in Rocks Clusters using eucalyptus Cloud computing" Proceedings of the International Conference on Management of Emergent Digital EcoSystems, pp. 61-66.
- B. Sotomayor, R. S. Montero, I. M. Llorente, e I. Foster. (2009) "Virtual infrastructure management in private and hybrid clouds". *IEEE Internet Computing*, 13(5):14-22.
- F.J. Fernandes, B. Schulze, e A.R. Mury, (2011) "Neblina – Espaços Virtuais de Trabalho para uso em Aplicações Científicas", Anais do XXIX Simpósio Brasileiro de Redes de Computadores e Sistemas Distribuídos, pp. 965-972.
- Science Clouds, (2013) "What are the Science Clouds?". Disponível em: <http://scienceclouds.org/> , Acessado em Julho.
- Napper, Jeffrey, and Paolo Bientinesi. (2009) "Can cloud computing reach the top500?." Proceedings of the combined workshops on UnConventional high performance computing workshop plus memory access workshop.

# Identificação dos fatores de maior impacto de aplicações Mapreduce

Fabiano da Guia Rocha

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso - *Campus Cáceres*  
Av. dos Ramires, s/n – 78.200-000 – Distrito Industrial – Cáceres – MT – Brasil

Programa de Pós-Graduação em Ciência da Computação  
Universidade Federal de São Carlos (UFSCar) – São Carlos, SP – Brasil

fabiano.rocha@cas.ifmt.edu.br

**Abstract.** *Mapreduce is a programming model, and a framework for the execution of applications which manipulate large data volumes in machines composed of thousands of processors/cores. Currently, Hadoop is the most widely adopted free implementation of Mapreduce. One of the main challenges in the scalability study of Mapreduce applications is the large number of configuration parameters of Hadoop. In this work we used the  $2^k$  factorial analysis evaluated the influence of a parameter or combination of parameters of a MapReduce application. In the tests different configurations of experiments allowed to estimate the effect of  $k$  factors (parameters) and the impact on performance and scalability of the application*

**Resumo.** *Mapreduce é um modelo de programação e um framework para a execução de aplicações que manipulam grandes volumes de dados em máquinas compostas por milhares de processadores ou núcleos. Atualmente, o Hadoop é a implementação como software livre de Mapreduce mais largamente adotada. Um dos principais desafios no estudo de escalabilidade de aplicações Mapreduce é o grande número de parâmetros de configuração da aplicação e do ambiente Hadoop. Neste trabalho utilizamos a análise  $2^k$  fatorial para avaliar a influencia de um parâmetro ou da combinação de um conjunto de parâmetros de uma aplicação Mapreduce. Nos testes diferentes configurações de experimentos permitiram avaliar o efeito de  $k$  fatores (parâmetros) e o impacto no desempenho e escalabilidade da aplicação.*

## Introdução

No modelo de programação Mapreduce, as tarefas de processamento são distribuídas entre os nós, implementando uma arquitetura *master-slave*, na qual existe um único mestre gerenciando um determinado número de escravos (*workers*). O nó mestre escalona as tarefas aos escravos, determinando qual nó irá realizar uma tarefa *map* ou uma tarefa *reduce*. Sempre que um nó escravo completar a execução de uma tarefa, ele sinaliza ao mestre, para que esse possa escalonar uma nova tarefa ao escravo que está disponível [DEAN and GHEMAWAT 2008].

Mapreduce é utilizado em conjunto com um sistema de arquivos distribuídos, especialmente projetado para dar suporte a aplicações distribuídas que manipulam

grandes volumes de dados (e.g. HDFS, Google GFS). White (2009) descreve que um sistema de arquivos distribuídos (*Distributed File System* - DFS) deve fornecer transparência de localidade e de redundância a fim de melhorar a disponibilidade dos dados e de funcionar como uma plataforma de armazenamento central.

No sistema que implementa o modelo Mapreduce, cada nó recebe blocos de dados (*chunks*) como entrada, os quais estão armazenados no sistema de arquivos do *cluster*. Então, cada nó processa em diferentes máquinas. O processamento das tarefas é abstraído em três fases distintas: a fase *map* e a fase *reduce*, acessíveis ao programador, e a fase intermediária chamada *shuffle* que é criada pelo sistema em tempo de execução [WHITE 2009].

Existem várias ferramentas que implementam o modelo de programação Mapreduce, cada qual criada para atender a necessidades específicas nas mais diversas linguagens de programação. As principais ferramentas são: Google Mapreduce (C++), Hadoop Mapreduce (Java), Greenplum (Python), Phoenix (C), Skynet (Ruby), MapSharp (C#), entre outras. A seguir, serão descritos o framework Hadoop e seus componentes.

Hadoop é a implementação mais conhecida e utilizada do modelo Mapreduce. O projeto foi desenvolvido tendo como foco o processamento em *clusters* com milhares de máquinas e sustentado sobre duas bases: Mapreduce, que provê as interfaces de programação, e HDFS (*Hadoop Distributed File System*), que fornece um sistema de arquivos distribuídos. HDFS se mostra um grande aliado ao projeto Hadoop, pois permite explorar a localidade de arquivos na arquitetura, escalonando tarefas aos nós que possuem os arquivos de entrada para execução [WHITE 2009].

Em ambientes distribuídos, Hadoop é capaz de gerenciar a arquitetura distribuída, realizando, de forma automática, atividades como o escalonamento de tarefas, a interação com o sistema de arquivos distribuídos e a troca de mensagens entre nós, permitindo que o usuário se preocupe apenas com a lógica da aplicação.

O ambiente Hadoop possui aproximadamente 190 parâmetros de configuração, dos quais aproximadamente 25 interferem mais significativamente no desempenho de um *job* [BABU 2010].

Nesse contexto, a fim de avaliar os limites de escalabilidade, faz-se necessário identificar os parâmetros de maior impacto, bem como encontrar a configuração ideal do conjunto de parâmetros que permitirá obter a máxima escalabilidade de uma aplicação. Portanto, foram criadas diferentes configurações de experimentos e cada uma delas é a combinação entre parâmetros e seus valores de referência. A análise dessa combinação tornou possível avaliar o seu impacto na execução.

O ambiente de simulação é composto por um conjunto de programas que criam os dados de entrada, executam as simulações e extraem as informações desejadas. Neste trabalho, o ambiente de simulação utilizado é o SimGrid, que simula plataformas como *clusters*, *grids* e sistemas P2P [CASANOVA, 2008].

Um arquivo *XML* permite descrever o poder computacional dos nós, a largura de banda dos *links* e a topologia da rede do sistema. Para simular as aplicações, optou-se pelo MRSG (*MapReduce Over SimGrid*), que simula aplicações Mapreduce sob o SimGrid [KOLBERG and ANJOS 2011]. O MRSG lê um arquivo texto (*mrsg.conf*) que

descreve os parâmetros da aplicação, tais como: o número de tarefas *map*, o tamanho do *chunk* de dados, o custo das tarefas *map* e *reduce*, entre outros.

## O Projeto de Experimentos $2^k$ Fatorial

Na análise da influência dos fatores na escalabilidade dos experimentos, optou-se pela abordagem do projeto  $2^k$  fatorial, com o objetivo de observar o impacto de cada um dos fatores no desempenho. Para o melhor entendimento, conceituam-se os 9 parâmetros utilizados nos experimentos:

- Número de nós: quantidade de nós que compõem o *cluster* simulado;
- Número de *reduces*: quantidade de tarefas *reduce* que serão executadas na simulação;
- *Chunk Size*: tamanho do bloco de dados em que a base deve ser particionada. Ex. base de dados de 1 Gigabyte particionada em 16 blocos de 64 Megabytes cada;
- Dfs Réplicas: quantidade de réplicas de cada bloco de dados na plataforma;
- *Map\_Output*: percentual de saída de dados produzidos pela fase *map*, considerando o valor da base de entrada. Por exemplo, para a base de dados de entrada de 1 Gigabyte, um *mapout* de 10% iria produzir 102,4 Megabytes de dados na fase *map*;
- *Map\_Cost* e *Reduce\_Cost*: custo de processamento por *byte* em cada fase. Esses dois parâmetros são utilizados para estimar a quantidade de trabalho a ser executado em cada tarefa *map* e *reduce*. Dessa maneira, para modelar corretamente a aplicação no simulador, faz-se necessário o ajuste desses custos de processamento.

Além dos sete parâmetros descritos, o simulador MRSG considera o número de *slots map* e *slots reduce*. Nessa análise, considera-se o número de *slots* igual ao número de *core* de cada nó da plataforma experimental, ou seja, o valor fixo de dois *slots map* e dois *slots reduce*.

Após a escolha dos parâmetros iniciais, desenvolveu-se um *shell script*, responsável por criar as  $n$  combinações de cada fator e nível e, para cada combinação possível, criar o arquivo de configuração que é utilizado como entrada do simulador MRSG. A execução das simulações foi automatizada por outro *shell script* que executa e extrai o *log* e as informações necessárias para a análise.

Os cálculos da análise  $2^k$  fatorial foram automatizados por meio da criação de uma planilha também utilizada para desenhar os gráficos que evidenciam os fatores de maior impacto.

## Fatores e seus níveis

O  $2^k$  fatorial considera o problema de analisar o impacto de 7 diferentes fatores em relação à variável de saída da aplicação ora simulada. Essa combinação de fatores e níveis foi utilizada como configuração da simulação no MRSG. Tais experimentos, considerando-se 7 fatores com 2 níveis cada, permitiram analisar 128 combinações

diferentes. Além das 128 combinações, foram geradas 32 combinações adicionais destinadas a representar o tempo de execução sequencial (considerando 1 nó) utilizado para o cálculo da eficiência.

A combinação de cada fator e de seus níveis foi gerada e salva em um arquivo, que juntamente com os arquivos de plataforma e de ambiente compõem os dados de entrada do simulador. Para todos os experimentos, a base de dados foi fixada em 20480 Megabytes (20 Gigabytes). Vale notar que o número de tarefas *map* lançadas para a execução é calculado pela divisão da base de dados pelo tamanho do *chunk*. Por exemplo, considerando a base de dados acima e o tamanho de *chunk* de 128 Megabytes, um total de 160 tarefas *map* (160 *chunks*) serão configuradas.

Após a execução da simulação, foram extraídos os seus *logs* com os dados necessários para a análise, em especial, o valor da eficiência. Esse valor foi calculado por meio do tempo da execução da configuração paralela e do tempo da execução da mesma combinação em uma plataforma sequencial. O valor de eficiência foi a variável de saída utilizada na análise do impacto dos fatores no experimento.

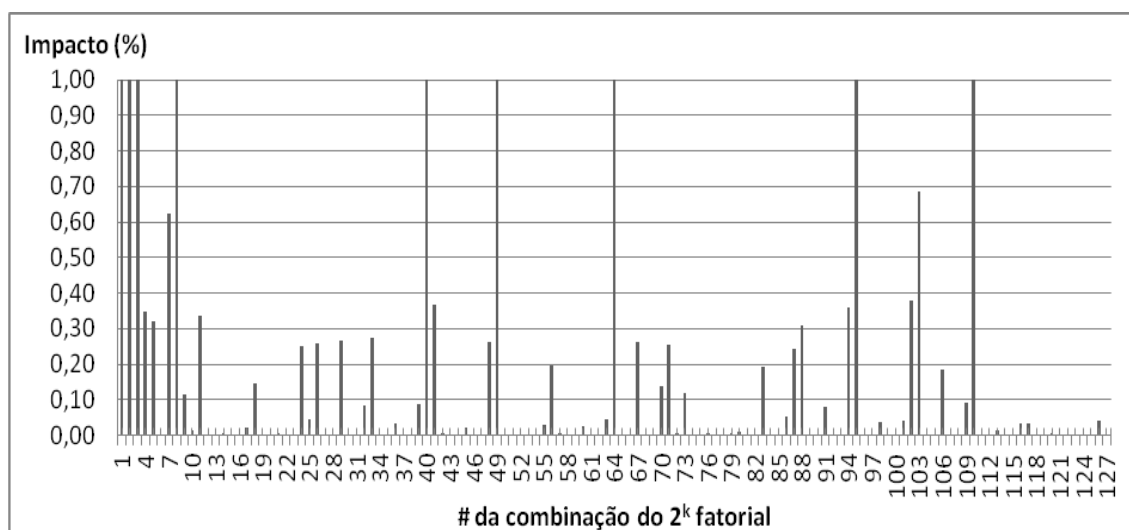
## Primeiro Experimento

Os fatores e os níveis analisados nesse primeiro experimento estão descritos na Tabela 1. A escolha desses níveis foi realizada com o objetivo de analisar, inicialmente, valores extremos de configuração e, nos experimentos seguintes, refinar esses valores.

**Tabela 1: Fatores e níveis do primeiro experimento 2<sup>k</sup> fatorial**

| Número de Nós | Número de Reduces | Chunk Size | Dfs réplicas | Map Output | Map Cost | Reduce_Cost |
|---------------|-------------------|------------|--------------|------------|----------|-------------|
| 10            | 20                | 16         | 1            | 1          | 1        | 1           |
| 1000          | 2000              | 128        | 10           | 1000       | 10000    | 10000       |

A representação gráfica (Figura 1) resultante do processamento das 128 configurações (1 a 127 no gráfico) possíveis entre fatores e níveis da tabela anterior descreve o impacto de cada combinação. Vale observar que, na grande maioria das combinações, o impacto foi inferior a 0,70%, o que se considera negligível.



**Figura 1: Percentual de impacto dos fatores para o primeiro experimento.**

Na Tabela 2, encontra-se o resultado da análise evidenciando os fatores e as combinações de fatores de maior impacto.

**Tabela 2: Percentual de impacto dos fatores sobre a eficiência do primeiro experimento.**

| Número do Experimento | Fator/Combinação de fatores                                | Impacto (%) |
|-----------------------|--|-------------|
| 02                    | Número de Nós  | 25,3        |
| 01                    | Número de <i>Reduces</i>                                   | 22,0        |
| 64                    | Número de <i>Reduces</i> + Map Output                      | 12,4        |
| 40                    | Número de <i>Reduces</i> + <i>Reduce_Cost</i>              | 10,8        |
| 110                   | <i>Reduce_Cost</i> + Map Output                            | 8,0         |
| 49                    | Número de <i>Reduces</i> + <i>Reduce_Cost</i> + Map Output | 4,6         |
| 03                    | <i>Reduce Cost</i>   | 4,2         |
| 95                    | Número de Nós + Map Output                                 | 3,4         |

Observa-se que, para a configuração analisada, o número de nós representa o fator de maior impacto. Em seguida, o número de tarefas *reduce* e sua combinação com a quantidade de dados produzida pela fase *map* (*map output*) refletem no desempenho da aplicação, como observado na Tabela 2 nos experimentos 01 e 64. Nos demais experimentos relacionados na tabela, observa-se que a combinação do número de tarefas *reduce*, o *map output* e o custo de processamento de tarefas *reduce* tiveram menor impacto.

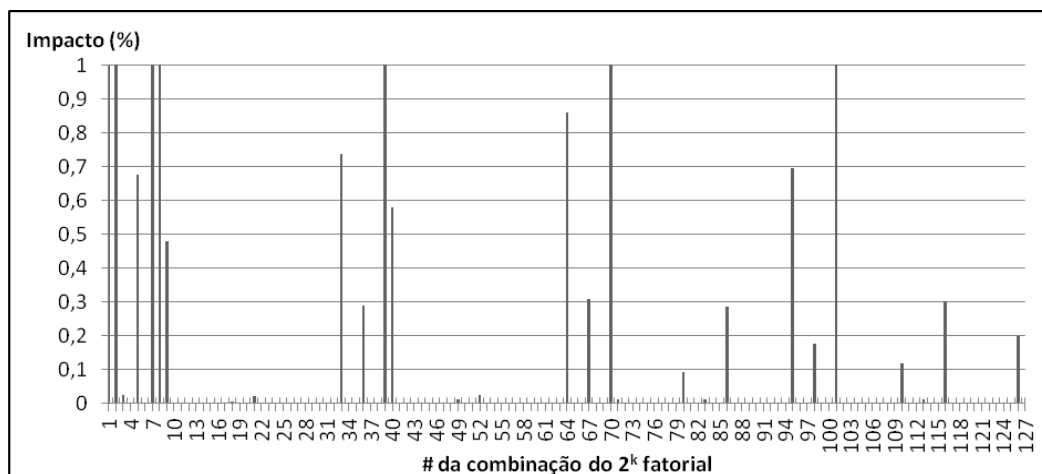
### Segundo Experimento

No segundo lote de experimentos, reduziram-se os valores dos níveis dos fatores para representar outro cenário de aplicação Mapreduce, conforme os valores descritos na Tabela 3. Com variações menores, tentou-se avaliar o impacto dos fatores em situações menos adversas.

**Tabela 3: Fatores e níveis do segundo experimento**

| Número de Nós | Número de <i>Reduces</i> | <i>Chunk Size</i> | Dfs réplicas | <i>Map Output</i> | <i>Map Cost</i> | <i>Reduce_Cost</i> |
|---------------|--------------------------|-------------------|--------------|-------------------|-----------------|--------------------|
| 10            | 20                       | 16                | 1            | 50                | 10              | 10                 |
| 1000          | 2000                     | 128               | 10           | 200               | 100             | 100                |

A simulação dessas 128 configurações possíveis produziu o resultado gráfico ilustrado pela Figura 2.



**Figura 2: Percentual de impacto dos fatores para o segundo experimento.**

Observa-se no gráfico (Figura 2) que, com a redução do percentual de saída de dados da fase *map*, o parâmetro *chunk size* é evidenciado como fator de alto impacto. Considerando que o *chunk size* refere-se à quantidade de dados que cada tarefa *map* deve processar, é natural observar que o número de nós seja um fator impactante, já que uma menor quantidade de nós reflete uma menor quantidade de tarefas *map* sendo executadas concorrentemente.

Na Tabela 4 encontra-se o resultado da análise, evidenciando os fatores e as combinações de fatores de maior impacto.

**Tabela 4: Percentual de impacto de cada fator sobre a eficiência do segundo experimento.**

| Número do Experimento | Fator/Combinação de fatores                                  | Impacto (%) |
|-----------------------|--|-------------|
| 02                    | Número de Nós  | 58,3        |
| 01                    | Número de <i>Reduces</i>                                     | 16,5        |
| 08                    | Número de <i>Reduces</i> + Número de Nós                     | 11,4        |
| 07                    | <i>Chunk Size</i>  | 3,0         |
| 39                    | Número de <i>Reduces</i> + Número de Nós + <i>Chunk Size</i> | 2,2         |
| 101                   | Número de Nós + <i>Chunk Size</i>                            | 1,4         |
| 02                    | Número de Nós  | 58,3        |
| 01                    | Número de <i>Reduces</i>                                     | 16,5        |

## Considerações Finais

Plataformas reais, tal como o Hadoop, possuem grande número de fatores que podem ser configuráveis e muitos deles causam impacto no desempenho das aplicações. A análise  $2^k$  fatorial permite identificar os fatores mais importantes na aplicação considerando uma variável de saída, que no caso do experimento realizado, utilizou-se a eficiência.

Nesta análise, os fatores mais impactantes foram o número de nós, o número de *reduces*, o tamanho do *chunk size*, a saída da fase *map* e o custo das tarefas *reduce*. Tais fatores serão utilizados nos experimentos de análise de escalabilidade descritos no próximo capítulo.

## Referências bibliográficas

- BABU, S. (2010) "**Towards Automatic Optimization of MapReduce Programs**", Duke University. Durham, North Carolina, USA.
- CASANOVA, H., LEGRAND, A. and QUINSON, M. (2008) "**SimGrid: a Generic Framework for Large-Scale Distributed Experiments**", 10<sup>th</sup> IEEE International Conference on Computer Modeling and Simulation.
- DEAN, J. and GHEMAWAT, S. (2008) "**Mapreduce: Simplified Data Processing on Large Clusters**", Communications of The ACM, Vol. 51, No. 1.
- KOLBERG, W. and ANJOS, J. C. S. (2011) "**MRSRG: a mapreduce simulator for desktop grids**", Instituto de Informática, UFRGS.
- WHITE, T. (2009) "**Hadoop: the Definitive Guide**", O'Reilly.

# Fragmentação de Memória em Sistemas Linux

Felipe Oliveira<sup>1</sup>, Daniel Costa<sup>1</sup>, Glaydson Vale<sup>1</sup>, David Beserra<sup>1</sup>, Jean Araujo<sup>1,2</sup>

<sup>1</sup>Unidade Acadêmica de Garanhuns - UAG  
Universidade Federal Rural de Pernambuco - UFRPE

<sup>2</sup>Centro de Informática - CIn  
Universidade Federal de Pernambuco – Recife – PE – Brasil

{flpdias14, danielalves1995cm, glaydson.gmv, dw.beserra}@gmail.com

jcta@cin.ufpe.br

**Abstract.** *This paper describes the use of Systemtap tool for search memory fragmentation occurrences in Linux, while accessing YouTube videos. Systemtap enables dynamic instrumentation of the Linux kernel, allowing multiple tests to be performed at runtime and user-friendly way. It was observed that the biggest cause of occurrence of fragmentation was the Flash Player*

**Resumo.** *Este artigo descreve o uso da ferramenta Systemtap na busca por ocorrências de fragmentação de memória em Linux durante o acesso a vídeos do YouTube. O Systemtap permite a instrumentalização dinâmica do kernel de sistemas Linux, possibilitando que vários testes sejam realizados em tempo de execução e de maneira amigável ao usuário. Foi observado que o maior causador de ocorrências de fragmentação foi o Flash Player.*

## 1. Introdução

Com o crescente número de aplicações disponíveis para os usuários, tornou-se comum se deparar com softwares que acarretam em problemas que envolvam uso indevido dos recursos computacionais. Dentre os problemas que se pode perceber, existem aqueles relacionados à má utilização de memória, que é um recurso escasso e valioso. O uso eficiente da memória impacta no desempenho de um sistema computacional. Durante a execução de um software, são realizadas várias alocações e liberações de espaços na memória. Com o decorrer destas operações, pequenos espaços de memória podem vir a se tornar inutilizáveis, o que é chamado de fragmentação de memória. A fragmentação da memória é um dos efeitos do envelhecimento de software, um fenômeno que provocando perdas significativas de desempenho, além de erros e travamentos do sistema [Grottke et al. 2008].

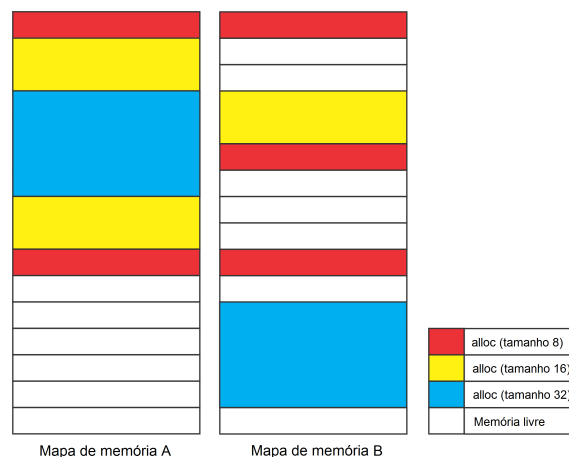
É difícil identificar os efeitos da fragmentação de memória, sendo fundamental monitorar constantemente o sistema. Devem-se utilizar técnicas que permitam monitorar os aspectos do sistema que se deseja observar, sem interferir em seu funcionamento [Lilja 2000]. Com o intuito de manter a disponibilidade e o bom funcionamento de tais sistemas essas práticas se mostram essenciais, tanto para otimizações de desempenho quanto para a detecção de erros.

Esse trabalho investigou as ocorrências de fragmentação de memória no sistema operacional Linux durante o acesso a *streams* de vídeo por um navegador web. Durante os experimentos foi utilizado o Systemtap [SystemTap 2014], uma ferramenta que permite a instrumentalização dinâmica de *kernels* Linux. Como resultados obtidos, são contabilizadas as ocorrências de fragmentação disparadas pelo *kernel*, mostrando a eficácia dessa estratégia.

O restante do artigo está organizado como segue: A Seção 2 aborda os conceitos básicos relacionados à fragmentação de memória. A Seção 3 apresenta a ferramenta SystemTap. A Seção 4 descreve os detalhes do experimento realizado, do ambiente de testes, e dos resultados obtidos. Encerrando o trabalho, a Seção 5 apresenta as conclusões e trabalhos futuros.

## 2. Fragmentação de memória

Em um computador, é o Sistema Operacional (SO) o responsável pelo uso eficiente dos recursos físicos [Wulf et al. 1974], sendo o *kernel* a porção do SO responsável pelo gerenciamento de processos e da utilização da memória [Hansen 1970]. Na memória principal utiliza-se duas formas de particionamento : dinâmico e fixo [Mamagkakis et al. 2006]. No primeiro caso, os processos alocam memória ao longo de sua execução. No segundo, a memória é dividida em áreas de tamanho fixo. Em ambas as formas podem ocorrer fragmentação de memória. A Figura 1 ilustra um exemplo de fragmentação de memória.



**Figura 1. Ilustração de mapa de memória - A) Memória contígua e B) Memória fragmentada**

Existem dois tipos de fragmentação de memória: interna e externa. A fragmentação interna ocorre quando há particionamento fixo. Neste tipo de particionamento, quando um processo solicita um espaço de memória para sua utilização, as chances que um processo tenha tamanho igual ao bloco alocado são pequenas. Assim, se um processo está utilizando apenas parte de um bloco, este ficará reservado para este processo não sendo permitido a sua utilização por outro.

A fragmentação externa ocorre no esquema de particionamento dinâmico. Neste esquema a memória será alocada e liberada por muitos processos, ocasionando espaços vazios de forma dispersa [Mamagkakis et al. 2006]. Se um processo necessita de espaço em memória, mesmo havendo memória disponível ela estará dispersa, o que faz com que seja procurado por espaço contíguo para a alocação.

Em razão de tais eventos ocorrerem a nível de *kernel*, monitorá-los não é uma atividade trivial. Por muitas vezes torna-se necessário a modificação do *kernel* para a implantação de algumas técnicas de depuração como *printk()*. O *printk()* é uma função semelhante ao *printf()* da linguagem C e utiliza macros para classificar as mensagens de acordo com sua importância [Pomerantz et al. 2009]. O uso da técnica consiste em alterar o código do *kernel*, colocando esta função nos pontos estratégicos que deseja-se analisar e exibindo mensagens a medida que

o *kernel* está sendo executado naquele ponto. Essa técnica demanda conhecimento detalhado do código fonte do *kernel* e não permitir novas medições enquanto o sistema está em execução. Uma alternativa mais simples a essas técnicas é o SystemTap, descrito na próxima seção.

### 3. SystemTap

O SystemTap é uma ferramenta de código aberto desenvolvida para monitorar as atividades do *kernel* de sistemas baseados em Linux, oferecendo informações detalhadas para análises de desempenho e detecção de erros[Eigler 2006]. Foi criada em uma parceria que incluiu empresas como IBM, Intel, Hitachi, Oracle e Red Hat [Jacob et al. 2009], [Prasad et al. 2005]. O SystemTap possui uma linguagem própria, similar a linguagem C. O usuário pode criar scripts que verifiquem eventos específicos do sistema, com segurança e possibilidade de reutilização dos testes em outras situações [Prasad et al. 2005]. Na maioria das distribuições Linux o SystemTap pode ser instalado através de repositórios.

#### 3.1. Funcionamento da ferramenta

O funcionamento do SystemTap consiste em rastrear os eventos denominados *probe points*, que podem ser a chamada ou o retorno de uma função, recebimento de pacotes de rede, ou mesmo contadores. Pode-se fazer *probe points* em qualquer função do *kernel* que não seja *inline* [Leitão 2010]. Quando o evento é disparado, chama-se o *handler* (manipulador) que coleta e mostra os dados desejados. Quando o manipulador termina sua execução, o *kernel* volta a sua rotina normal [Prasad et al. 2005] [Jacob et al. 2009]. O fluxo de execução do SystemTap pode ser visto na Figura 2.

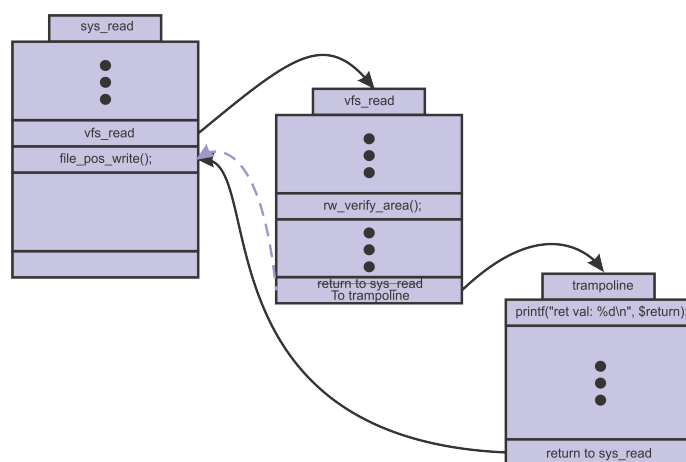


Figura 2. Fluxo de execução do SystemTap - adaptado de [Jacob et al. 2009]

A Figura 3 exemplifica a lógica de execução do SystemTap. O código passa por um processo de elaboração, onde são verificadas as referências a outras bibliotecas; depois é traduzido para a linguagem C e logo em seguida é compilado como um módulo do *kernel*. Posteriormente, este módulo é carregado no *kernel*, dando início a coleta dos dados. Após o período de coleta, o módulo é descarregado e os dados são enviados para a saída padrão [Prasad et al. 2005].

O SystemTap utiliza uma linguagem simples e que se assemelha à linguagem C. Segue abaixo o exemplo de um script do SystemTap:

```
# stap -ve 'probe begin { log("hello world") exit () }'
```

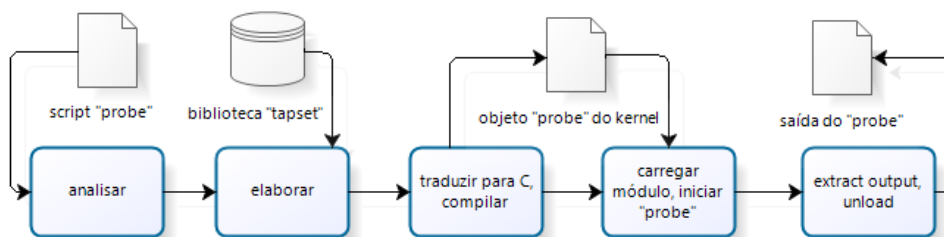


Figura 3. Passos de execução do SystemTap - adaptado de [Prasad et al. 2005]

## 4. Estudo experimental

Para verificar a eficácia da ferramenta em monitorar o *kernel* do Linux, foi realizado um experimento onde buscou-se detectar eventos de fragmentação de memória. Foi realizada uma simulação do uso de um navegador de internet, onde aplicou-se uma carga de trabalho para simular as requisições do usuário.

De fato, a estratégia de monitoramento não se restringe somente às aplicações web, podendo ser aplicada em qualquer outro tipo de sistema, como editores de texto, jogos, etc. Contudo, a escolha do cenário de testes se deu em razão de que os recursos multimídia são muito utilizados na Internet, afetando, portanto, milhões de usuários.

### 4.1. Ambiente de testes

Foi utilizado no experimento um computador com processador AMD Athlon II X2 3.00 GHz e 2 GB de memória principal. O SO utilizado foi o CentOS 6.5, *kernel* 2.6.32-431.el6. O navegador utilizado foi o Mozilla Firefox versão 24.2.0 junto com o Adobe Flash Player 11.2.202.332.

A carga de trabalho foi automatizada com o conjunto de ferramentas Selenium IDE [Selenium 2014] para testes funcionais, extensão do *Firefox* que simula ações do usuário como cliques e digitação de textos na página testada. As ações que serão executadas são geradas para uma linguagem nativa, mas o script que também pode ser exportado para linguagens de programação como Ruby, C#, Java ou Python. O script gerado para o teste foi exportado para a linguagem Python para facilitar a execução do experimento. O arquivo que inicializa o teste é executado uma única vez quando o monitoramento é iniciado.

A Figura 4 ilustra o *workload* utilizado no experimento, que simulou o acesso e a visualização de vídeos do YouTube [Google 2014]. Inicialmente realizou-se uma pesquisa de vídeos utilizando a palavra-chave “Cold Play”. O primeiro resultado foi selecionado e, após 30 segundos, foi selecionado outro vídeo (clcando e visualizando) até completar um ciclo de 5 visualizações. Ao fim de cada ciclo, retornava-se à página inicial do *site*. Aguardou-se mais trinta segundos e um novo ciclo foi iniciado. O intervalo de trinta segundos entre dois ciclos serve para garantir que a última instrução do *script* fosse executada. O *workload* foi executado indefinidamente até o final do experimento quando encerrou-se o monitoramento.

Para verificar a ocorrência de fragmentação monitorou-se a chamada *mm\_page\_alloc\_extfrag* que se refere a fragmentação de memória. O *script* utilizado monitora quando este evento ocorre e incrementa uma variável sempre que o a quantidade de memória liberada for menor que a que está alocada.

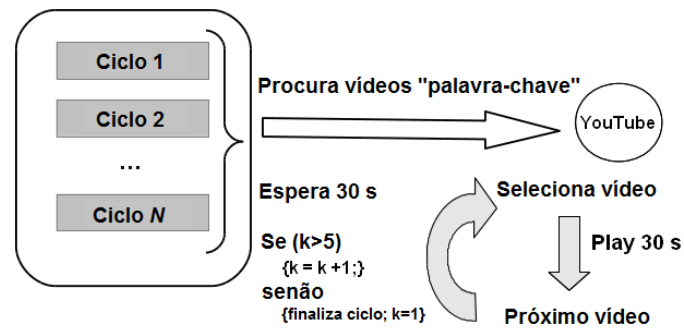


Figura 4. Carga de trabalho gerada com o Selenium IDE

## 4.2. Resultados

O script foi executado durante 4 horas. Os resultados obtidos estão exibidos na Figura 5. Nela são mostradas quais as aplicações geraram mais fragmentação. Durante a execução do experimento os processos que dispararam ocorrências de fragmentação foram o *plugin-containe*, o *Xorg*, *kthreadd*, e o *Firefox*.

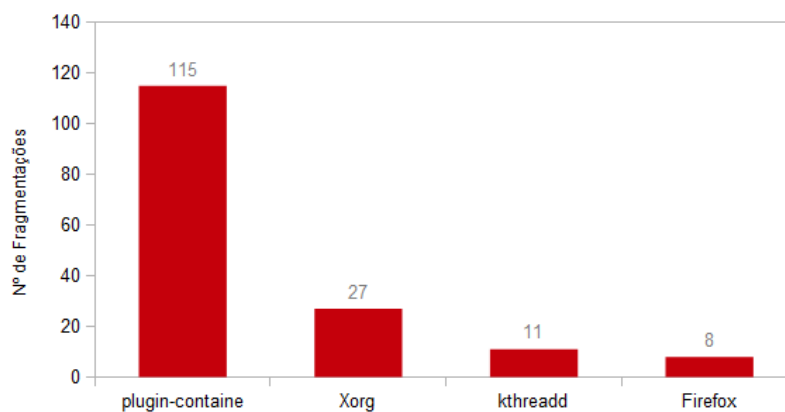


Figura 5. Ocorrências de fragmentação

O processo *plugin-containe* foi o que mais apresentou ocorrências de fragmentação, totalizando 115 eventos. Ele é o responsável por carregar os *plugins* que o *Firefox* utiliza para exibir conteúdo em formatos que o navegador não é projetado para exibir [Mozilla 2014]. Dentre esses *plugins*, encontra-se o *plugin Flash Player*, responsável pela exibição dos vídeos do *YouTube*.

O segundo processo em ocorrências de fragmentação de memória é o *Xorg*, com 27 ocorrências. O *Xorg* é uma implementação pública e software livre do *X11 X Window System*, que fornece a interface gráfica para o SO. O processo *kthreadd*, pai de todas as *threads* do *kernel*, apresentou 11 eventos de fragmentação. Já o processo do navegador *Firefox* foi o responsável por somente 8 ocorrências de fragmentação.

Diante da identificação de qual processo é o mais afetado pelo envelhecimento no ambiente testado, pode-se propor estratégias de rejuvenescimento de software (fenômeno contrário ao envelhecimento) para mitigar os efeitos causados pelo envelhecimento. Como o rejuvenescimento de software se trata de uma ação específica para cada ambiente, abrimos

então um leque de possibilidades futuras que possam identificar qual estratégia seria a mais indicada para o cenário investigado.

## 5. Conclusões e trabalhos futuros

Neste artigo foi exposto como uma instrumentalização dinâmica do *kernel* pode ser útil para monitorar os aspectos intrínsecos de seu funcionamento. Viu-se que o SystemTap é uma ferramenta de instrumentalização que permite obter informações do que está ocorrendo no nível do *kernel*, sem a necessidade de alterações em seu código fonte. Sua aplicação na análise da presença de fragmentação de memória no acesso de *streams* de vídeo na internet auxiliou a descobrir que o processo/aplicativo com menor quantidade de ocorrências de fragmentação é o navegador Firefox, seguido pelos processos internos do sistema operacional.

Como trabalhos futuros, pretendemos propor estratégias de rejuvenescimento específicas para o cenário investigado nesse artigo.

## Referências

- Eigler, F. C. (2006). Problem solving with systemtap. *Proceedings of the Linux Symposium*, 1:261–268.
- Google (2014). YouTube, LLC. Acesso em 29/02/2014. Disponível em: <http://www.youtube.com/>.
- Grottke, M., Matias, R., and Trivedi, K. (2008). The fundamentals of software aging. In *Proc of 1st Int. Workshop on Software Aging and Rejuvenation (WoSAR), in conjunction with 19th IEEE Int. Symp. on Software Reliability Engineering*, Seattle.
- Hansen, P. B. (1970). The nucleus of a multiprogramming system. *Commun. ACM*, 13(4):238–241.
- Jacob, B., Larson, P., Leitão, B. H., and Silva, S. A. M. M. (2009). *Systemtap: Instrumenting the Linux kernel for analyzing performance and functional problems*. IBM Redbook.
- Leitão, B. (2010). Depuração de kernel com systemtap. *Linux Magazine*, 1(67):58–64.
- Lilja, D. J. (2000). *Measuring Computer Performance: A Practitioner's Guide*. Cambridge University Press, New York.
- Mamagkakis, S., Baloukas, C., Atienza, D., Catthoor, F., Soudris, D., and Thanailakis, A. (2006). Reducing memory fragmentation in network applications with dynamic memory allocators optimized for performance. *Comput. Commun.*, 29(13-14):2612–2620.
- Mozilla (2014). Mozilla firefox. Acesso em 21/03/2014. Disponível em: <http://www.mozilla.org/pt-BR/firefox/new/>.
- Pomerantz, O., Burian, M., and Salzman, P. J. (2009). *The Linux Kernel Module Programming Guide*. Quicksmart University Guides Series. CreateSpace.
- Prasad, V., Cohen, W., C., E. F., Hunt, M., Keniston, J., and Chen, B. (2005). Locating system problems using dynamic instrumentation. In *Linux Symposium*.
- Selenium (2014). Selenium IDE. Acesso em 01/03/2014. Disponível em: <http://www.seleniumhq.org/projects/ide/>.
- SystemTap (2014). Systemtap. RedHat, Inc. Acesso em 11/03/2014. Disponível em: <http://sourceware.org/systemtap>.
- Wulf, W., Cohen, E., Corwin, W., Jones, A., Levin, R., Pierson, C., and Pollack, F. (1974). Hydra: The kernel of a multiprocessor operating system. *Commun. ACM*, 17(6):337–345.

## Simulador do Processo de Instalação do Rocks Clusters com uso de Técnicas de IHC

David Beserra<sup>1</sup>, Anderson Cavalcanti<sup>1</sup>, Luciano Souza<sup>1</sup>, Rubens Karman<sup>2</sup>, Jean Araujo<sup>1</sup>, Sergio Mendonça<sup>1</sup>

<sup>1</sup>Unidade Acadêmica de Garanhuns – Universidade Federal Rural de Pernambuco  
Garanhuns – PE – Brasil

<sup>2</sup>Escola de Computação Inteligente – Universidade de Pernambuco  
Recife – PE – Brasil

**Abstract.** *Parallel programming courses needs optimize the process of configuring the cluster structures employed, because it is not the primary focus of these courses. In order to assist in solving this problem is presented in this work a simulator of the Rocks Clusters installation process. The simulator uses principles of Human-Computer Interaction to prevent mistakes are made by students during the process.*

**Resumo.** *Disciplinas sobre programação paralela necessitam otimizar o processo de configuração das estruturas de cluster que utilizam, por não ser esse seu foco primário. Para auxiliar na resolução desse problema é apresentado neste trabalho um simulador do processo de instalação do Rocks Clusters. O simulador utiliza princípios de Interação Humano-Computador para impedir que erros sejam cometidos por estudantes durante o processo.*

### 1. Introdução

Em diversos cursos de Ciência da Computação, tópicos relacionados à programação paralela e computação em cluster são oferecidos em disciplinas como Sistemas Distribuídos, Arquitetura de Computadores e em optativas específicas de programação paralela. Na disciplinas cujo enfoque principal é a programação paralela e não as infraestruturas de computação em si, torna-se necessário diminuir o tempo gasto com o ensino e implementação dessas infraestruturas. Dentre elas, o principal tipo são os clusters beowulf de alto desempenho para execução de aplicações que utilizem a Interface de Passagem de Mensagens (MPI) [Mello et al. 2010] .

O treinamento prático na montagem e instalação de clusters esbarra em algumas dificuldades. A primeira delas situa-se no âmbito infraestrutural, pois, para que estudantes, individualmente ou em pequenos grupos, possam praticar a instalação de clusters é necessário prover uma quantidade mínima de dois computadores por estudante/grupo. Essa é a quantidade mínima de elementos necessários em um cluster. A segunda dificuldade está no processo de instalação dos softwares necessários para o cluster, que demanda a instalação de vários pacotes específicos de software em conjunto com uma distribuição do Sistema Operacional (SO) Linux.

Para simplificar o processo de instalação de clusters, sistemas focados na facilidade de instalação foram desenvolvidos, como o Rocks Clusters, SO baseado em

CentOS Linux [Papadopoulos, Katz e Bruno. 2003]. Entretanto, mesmo que o Rocks Clusters seja orientado à simplicidade, ainda requer treinamento para sua operação. Um inconveniente para uso desse sistema em ambiente letivo é a necessidade de ter computadores com alto poder de processamento, que suportem sua instalação. Outro inconveniente é o tempo de instalação, que é de aproximadamente uma hora em um computador com dois núcleos de processamento e 2 GB de memória principal. Proporcionar computadores com poder computacional adequado, e em quantidade suficiente para que cada aluno ou equipe de alunos implemente seus próprios clusters com o Rocks Clusters é bastante oneroso, sendo inviável para instituições com poucos recursos financeiros.

Os autores observaram que o instalador do Rocks, embora simples de usar, apresenta algumas ambiguidades que causam dúvidas durante o processo de instalação. Essas ambiguidades podem levar a dúvidas, e conseqüentemente a erros durante a instalação, atrasando o processo de aprendizado. O uso de princípios de Interação Humano-Computador, como os relacionados à prevenção e recuperação de falhas auxiliam a impedir ou minimizar os erros durante um processo interativo, como o da instalação do Rocks Clusters.

Neste trabalho é apresentada uma ferramenta que simula o processo de instalação do Rocks Clusters. Os objetivos dessa ferramenta são reduzir o tempo necessário para o treinamento na instalação do Rocks Clusters, bem como reduzir a quantidade de recursos necessários para tal. Para facilitar o aprendizado e impedir que estudantes cometam erros por falta de conhecimento e de ambiguidades presentes no processo de instalação do Rocks, foram utilizados princípios de IHC no desenvolvimento do simulador. Este artigo está dividido em 5 seções: a seção 2 apresenta os trabalhos relacionados, enquanto a seção 3 clarifica os conceitos de IHC necessários para o entendimento deste trabalho. A seção 4 descreve o simulador proposto, mostrando onde a IHC foi utilizada. A seção 5 discute o trabalho e a seção 6 encerra o trabalho com as considerações finais.

## 2. Trabalhos Relacionados

Algumas alternativas para auxiliar no processo de instalação de clusters estão disponíveis na literatura. Basicamente, essas alternativas podem ser divididas em dois tipos: (i) clusters de máquinas virtuais instalados sob a infraestrutura de laboratórios de ensino já estabelecidos, como em [Johnson et al. 2011] e em [Beserra et al. 2013], (que implementa um cluster virtualizado com o VirtualBox e utilizando o Rocks Clusters), e (ii) clusters de máquinas virtuais hospedados em nuvens computacionais, como em [Ivica et al. 2009]. Essas alternativas possibilitam a economia com a quantidade de elementos computacionais necessários, porém ainda necessitam de computadores de alto desempenho. O tempo para uma única tentativa de instalação permanece praticamente o mesmo que em um ambiente não virtualizado, e não impedem erros dos usuários. Embora os erros durante a instalação não comprometam o desempenho dos hospedeiros, colaboram para aumentar o custo temporal por tentativa.

Uma alternativa ao uso de máquinas virtuais é a simulação dos processos de instalação, que requer poucos recursos e permite reproduzir procedimentos com grande

precisão. Em [Medeiros et al 2011] temos um exemplo bem sucedido de uso de simulação para ensino de processos de E/S em Sistemas Operacionais. A simulação, quando utiliza técnicas de IHC, pode ser mais eficaz em seu propósito educacional, uma vez que IHC pode colaborar para a melhora dos processos de interação de qualquer sistema de software interativo [Barbosa e Silva. 2010].

### 3. Fundamentação Teórica em IHC

Interação humano-computador é o estudo da interação entre pessoas e computadores [Barbosa et al 2010]. Neste trabalho, as principais técnicas de IHC utilizadas foram o conceito de comunicabilidade, os princípios de Gestalt e as técnicas de *design* orientado a prevenção de erros e rupturas comunicativas. A **comunicabilidade** é a capacidade de um sistema comunicar aos seus usuários quais as intenções de seus projetistas ao conceber o sistema. Os princípios de Gestalt explicam como o ser humano percebe padrões. Destes, foram utilizados neste trabalho os princípios de **proximidade** (objetos próximos são reconhecidos como um só), **boa continuidade** (traços contínuos são percebidos mais rapidamente que contornos aleatórios) e **região comum** (objetos dentro de uma região confinada são percebidos como um grupo). O mecanismo de prevenção de erros utilizado foi o de **prevenção ativa**, que consiste em impedir que o usuário use de processos de interação inadequados.

### 4. Learn Rocks Easy

O Learn Rocks Easy<sup>1</sup> é um simulador WEB do processo de instalação do Rocks Clusters que simula todos os passos utilizados na instalação, reproduzindo todas as telas de instalação com algumas modificações baseadas em técnicas de IHC. A introdução de IHC objetiva impedir que o usuário do sistema cometa erros durante a instalação.

Na tela inicial, mostrada na figura 1, existe um botão “Experimente”, que ao receber um clique inicia o processo de simulação. Entretanto, se o usuário for inexperiente ao ponto de não conhecer os conceitos mais elementares sobre computação de alto desempenho, computação em cluster e Rocks Clusters; pode obter informações ao clicar em *links* disponibilizados, que redirecionam para os respectivos verbetes na Wikipédia. Nesta tela foram utilizados alguns dos princípios de Gestalt, como o de região comum, proximidade e continuidade, que estão assinalados.



Figura 1. Tela inicial do Learn Rocks Easy, técnicas de IHC utilizados em vermelho.

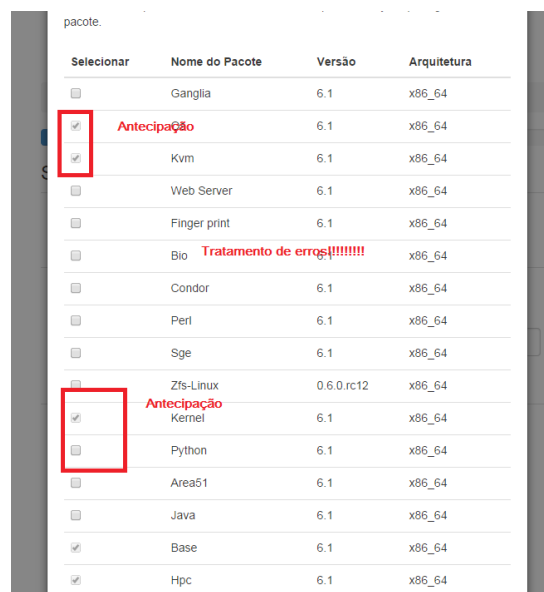
1 <http://agenciaimax.com/ihc/>.

Ao optar por experimentar a ferramenta, o usuário depara-se com a tela exibida na figura 2. Esta tela apresenta duas opções de origem para obter os pacotes de instalação, local, através de um CD/DVD ou pela Internet, através de um servidor específico para tal. Nesta tela, aparece uma barra de progressão que indica em que estado do processo de instalação o usuário se encontra. Esta barra não está presente no instalador original do Rocks, sendo uma inovação que utiliza do princípio da comunicabilidade para situar o usuário no espaço-tempo do problema.



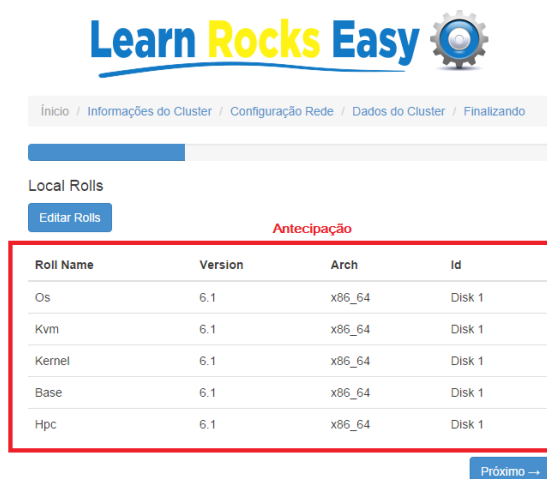
**Figura 2. Início da simulação de instalação.**

Quando o usuário seleciona uma fonte de onde os pacotes de instalação serão obtidos (ao clicar nos botões “CD/DVD-Local” ou “Baixar”, na tela 2), o simulador apresenta um menu suspenso com uma lista de pacotes a serem selecionados para a instalação, conforme exibido na figura 3. Alguns pacotes de instalação obrigatória já vem selecionados, como forma de proporcionar antecipação a possíveis erros do usuário. No instalador original do Rocks Clusters, esses pacotes não vem pré-selecionados, o que pode levar um usuário a cometer um erro fatal na instalação. Com o uso dessa técnica de IHC, este erro será evitado, economizando tempo.



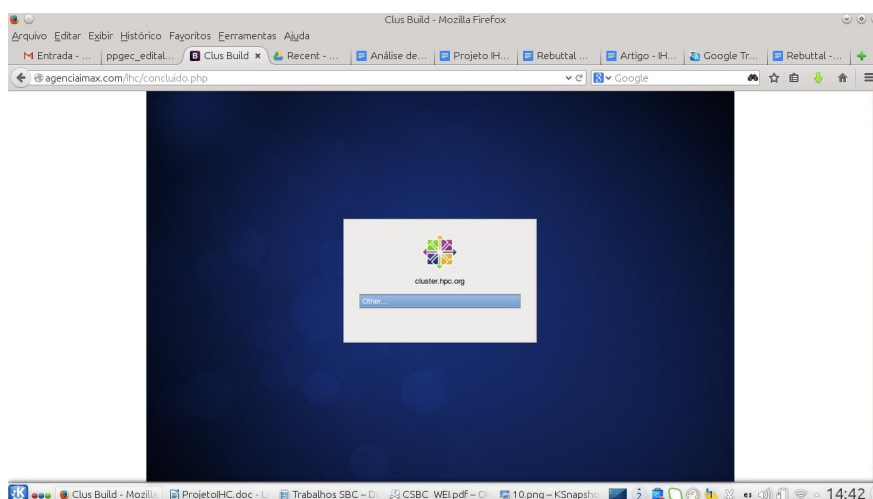
**Figura 3. Menu suspenso para seleção dos pacotes que serão instalados**

Ao selecionar os pacotes que devem ser instalados, o simulador passa a uma tela, que mostra quais pacotes foram selecionados, antes de prosseguir com a instalação. O usuário tem a opção de voltar atrás e modificar os pacotes, bem como refazer todas as ações, caso tenha dúvidas. Esta ação é exibida na figura 4. Todas as telas do simulador tem comportamentos parecidos.



**Figura 4. Pacotes selecionados para instalação.**

Após a conclusão de todos os passos de instalação, o sistema apresenta um comportamento tal qual exibido na figura 5, que representa a tela inicial pós instalação do Rocks Clusters. Esta tela indica o fim da instalação.



**Figura 5. Sistema instalado com sucesso**

## 5. Discussão e Considerações Finais

O uso dessa ferramenta de simulação é mais adequado para o treino na instalação de clusters com o Rocks Cluster do que outras soluções, como os clusters de VMs. Alguns dos benefícios identificados no uso dessa ferramenta são:

- Mobilidade no treinamento – Por ser uma ferramenta WEB, existe a possibilidade do usuário realizar seu treinamento remotamente, em qualquer lugar e horário que desejar;
- Redução de custos – O simulador utiliza poucos recursos, podendo ser executado em computadores com menos recursos e portanto, mais baratos. Além disso, basta um computador por aluno;
- Versatilidade - Possibilidade de ser empregue em conjunto com outras soluções, como os clusters de VMs. Esses clusters podem ser utilizados em paralelo ao uso do simulador ou depois, conforme um professor julgue oportuno;
- Economia de tempo e segurança no aprendizado – Com a utilização de técnicas de IHC que impedem que o usuário cometa erros em conjunção com técnicas que auxiliam na aprendizagem, o usuário não erra na simulação e aprende de uma única vez, economizando tempo e aprendendo eficazmente.

## 6. Considerações Finais e Trabalhos Futuros

O uso de simulação de processos interativos permite a economia de recursos físicos e temporais, além de proporcionar segurança e solidez no aprendizado, quando utilizada em conjunto com técnicas de IHC. Como trabalho futuro, os autores pretendem coletar dados de uso da ferramenta, para traçar um perfil de seus usuários e elaborar um processo de iteração ainda mais elaborado.

## Referências

- Beserra, D.W.S.C., Borba, A., Souto, S.C.R.A., Melo, C, Sousa. Y, de Andrade, M.J.P., de Araújo, A.E.P. e Souza. E (2013) "Ambiente Virtualizado para Ensino de Programação Paralela e Computação em Cluster." Em: *Workshop de Educação em Informática - 2013*. SBC, Maceió.
- Johnson, E., Garrity, P., Yates, T., e Brown, R. (2011) "Performance of a Virtual Cluster in a General-purpose Teaching Laboratory", In: *2011 IEEE International Conference on Cluster Computing*. Pp. 600-604. IEEE.
- Mello, T. C. Schulze, B. Pinto, R. C. G. e Mury, A. R. (2010) "Uma análise de recursos virtualizados em ambiente de HPC", Em: *Anais VIII Workshop em Clouds, Grids e Aplicações, XXVIII SBRC/ VIII WCGA*, SBC, Gramado, pp. 17-30.
- Papadopoulos, P. M., Katz, M. J., e Bruno, G. (2003). NPACI Rocks: Tools and techniques for easily deploying manageable linux clusters. *Concurrency and Computation: Practice and Experience*, 15(7-8), 707-725.
- Ivica, C. Riley, J.T. e Shubert, C. (2009) "StarHPC - Teaching parallel programming within elastic compute cloud", In: *Proceedings of the International Conference on Information Technology Interfaces, ITI*, pp. 353-356.
- Medeiros, T. R. Souza, C. C., Sousa, T. D. N. Gadelha, R. N. S. Silva, E. L. Júnior J. B. D., (2011) "IO Simulator: Um Simulador de Dispositivos de Entrada e Saída para Auxiliar o Ensino de Sistemas Operacionais", In: *Anais do WEI*, SBC, Natal-BR.
- Barbosa, S. D. J. e Bruno, S.S. (2010) *Interação humano-computador*. Elsevier, 2010.

# Metodologia de Gerenciamento para Projetos inovadores com interação Público-Privado

Rodrigo de Sousa e Silva, Nilton Hideki Takagi, Renato Júnior Caetano dos Santos, Josiel M. Figueiredo

Instituto de Computação - Universidade Federal de Mato Grosso (UFMT)

Av. Fernando Corrêa da Costa, nº 2367 - Bairro Boa Esperança. Cuiabá - MT - 78060-900

rodrigo.ross.comp@gmail.com, nilton@ic.ufmt.br,  
renatos.caetano@yahoo.com.br, josiel@ic.ufmt.br

**Abstract.** *On research environments the uncertainties due to the many variables that influence the outcome accompany the growth of information and data produced on the planet. Many projects happen in parallel, some with benefits in common, and there are variables that can be controlled and depend on the expertise of stakeholders, of the scenarios involved and conducted management fashion. To help researchers manage their projects and view the interrelation with the other, we propose a methodology that will guide the phases since the capture of the demand until the closure of the project.*

## **Keywords**

*Methodology of Project Management; PMO; Management of Research Projects*

**Resumo.** *Nos ambientes de pesquisa as incertezas em virtude da quantidade de variáveis que influenciam no resultado acompanham o crescimento de informações e dados produzidos no planeta. Muitos projetos acontecem em paralelo, algumas com benefícios em comum, e existem variáveis que podem ser controladas e dependem da expertise dos stakeholders, dos cenários envolvidos e da forma de gerenciamento realizado. Para auxiliar os pesquisadores a gerenciar seus projetos e ver a inter-relação com os demais, propomos uma metodologia que irá nortear as fases desde a captação da demanda até o encerramento do projeto.*

## **Palavras-chave**

*Metodologia de Gerenciamento de Projetos; PMO; Gerenciamento de Projetos de Pesquisa*

## **1. Introdução**

Um grupo de pesquisa é formado por alunos e professores com o objetivo de pesquisar ou gerar determinado conhecimento em uma ou mais linhas de pesquisas, com base em um ou mais projetos (UFES, 2014). Esses envolvidos acabam compartilhando ferramentas e o espaço das instituições participantes. Como atualmente as informações crescem de forma exponencial e os grupos não de pesquisa não seguem um padrão, estão sujeitos a vários imprevistos decorrentes das diversas variáveis e pela falta de

planejamento. Para minimizar esse tipo de impacto, as técnicas de gerenciamento de projeto auxiliam no controle podendo maximizar os resultados da pesquisa.

Kerzner (2006) afirma que há várias vantagens no uso de gerenciamento de projetos, estes são: aumento da chance de sucesso do projeto devido a uma análise mais aprofundada dos riscos dessa forma mantendo sempre que possível contramedidas preparadas; evitar surpresas durante a execução das atividades o que pode ocorrer muito facilmente, nem sempre as coisas saem como o planejado e uma melhor preparação visa manter de antemão essas surpresas mais previsíveis; facilitar e orientar as revisões do projeto mantendo sempre documentos arquivados sobre o projeto o que torna mais fácil revisá-lo e alterá-lo visto que todos os procedimentos estarão de alguma forma documentados; otimização e alocação de recursos; documenta e facilita estimativas para novos projetos; agiliza as tomadas de decisão, devido ao maior controle sobre os recursos e a possibilidade de obter dados mais rápidas e processá-lo para obter informações mais precisas; e por fim um amadurecimento em suas metodologias de trabalho.

O ambiente de trabalho em questão é no Instituto de Computação da Universidade Federal do Estado de Mato Grosso, no qual há vários grupos de pesquisa, com diferentes líderes e focos de trabalho. A proposta desse trabalho inicial é criar uma metodologia que será base para o Escritório de Projetos para auxiliar os grupos a maximizarem os seus esforços no *core* do conhecimento de cada área envolvida.

## 2. O Project Management Office

O PMO (*Project Management Office*), ou também conhecido em português como Escritório de projetos ou Escritório de Gerenciamento de projetos é uma unidade organizada para que se possa gerenciar projetos, programas e portfólios. Essa estrutura seria o meio de comunicação entre o planejamento estratégico da empresa e a execução dos projetos, permitindo assim tomar as melhores medidas para o andamento do projeto.

### 2.1 Níveis da organização para implantar o PMO

Crawford (2001) classificou o PMO em 3 níveis de acordo com algumas funções desempenhadas e com posicionamento organizacional.

**Tabela 1 – Níveis de PMO**

| Nível   | Foco                                  | Detalhes   |
|---------|---------------------------------------|--|
| Nível 1 | Em projetos e/ou programa individuais | Este PMO é geralmente aplicado em organizações com estrutura organizacional onde os membros da equipe tendem a ter mais contato com o gerente funcional do que o gerente do projeto  |
| Nível 2 | Em processos e otimização de recursos | Neste nível, o foco principal do PMO deixa de ser o de controlar e contornar problemas passando a ser mais metodológico, se responsabilizando também para compartilhar uma cultura de melhor gerência de projetos e no controle da alocação de recursos, buscando evitar o desbalanceamento. |
| Nível 3 | É na estratégia, visa à organização   | O PMO funciona durante toda a escala da organização espalhando seu conhecimento e forma de   |

|  |              |  |
|--|--------------|--|
|  | como um todo | trabalho, assim mesmo que o departamento será mais autônomo e o PMO servirá como forma a aumentar a integração com as demais áreas de negócio, e acabar com possíveis conflitos. |
|--|--------------|--|

O *Project Management Office* do Instituto de Computação (PMO-IC) da UFMT estará alinhado ao Nível 2 proposto, considerando o desenvolvimento da metodologia com apoio aos grupos de pesquisa, fornecendo recursos materiais e humanos para auxiliá-los na gestão dos projetos.

## 2.2 Implantando o PMO

A análise da organização é o alicerce para definição dos conceitos do escritório de projeto, caracterizando o funcionamento de acordo com a especificidade do ambiente (DINSMORE, 1990). Um passo inicial para a implantação dos serviços a serem prestados pelo escritório de projetos é a obtenção do acordo com o Sponsor da instituição sobre os mesmos. Cleland (2002) indica primeiramente definir o escopo do trabalho, as competências e as funções dos membros do escritório de projeto, pois eles determinarão a quantidade de apoio que pode ser fornecido pelo PMO. Neste momento seria essencial elaborar um documento com as informações básicas do projeto de implantação do PMO, contendo o objetivo para se estabelecer o escritório de projetos, as hipóteses e restrições associadas ao projeto, e os resultados esperados, como forma de manter controle sobre o andamento do projeto (DINSMORE, 1999).

Kerzner (2006) também assegura que a maturidade em gerenciamento de projetos pode ser definida como “o desenvolvimento de sistemas e processos que são por natureza repetitivos e garantem uma alta probabilidade de que cada uma delas seja um sucesso”. Contudo, esses sistemas e processos repetitivos não garantem o sucesso, mas apenas aumentam a probabilidade de sua ocorrência. Então vemos que o gerenciamento de projetos fornece o caminho e os meios de se obter o sucesso na gerencia.

## 3. A Metodologia do PMO-IC

Considerando as variáveis dos projetos de pesquisa envolvidos no PMO-IC, foi proposto uma metodologia contendo 5 fases: Captação, Planejamento, Execução, Monitoramento e Encerramento. As fases são descritas na sequência:

### 3.1 Fase de Captação

Essa fase se inicia com a empresa enviando sua demanda de negócio para a instituição. A partir daí o PMO-IC fica encarregado de cadastrar essa demanda, o que significa que ele fará uma análise sobre a mesma, identificando, justificando sua necessidade e definindo seus resultados esperados e seu alinhamento ao planejamento estratégico. A demanda é divulgada para os grupos de pesquisa e aqueles que demonstrarem interesse participam de uma reunião organizada pelo PMO-IC no qual há uma exposição da demanda e a seleção do gerente de projeto. Após esses momentos, entra na fase de planejamento, no qual o gerente de projetos deve iniciar o Termo de Abertura do Projeto (TAP).

### 3.2 Fase de Planejamento

O patrocinador deve validar o TAP, plano de trabalho e o orçamento para seguir no planejamento do projeto. Conforme a figura 1, os processos do planejamento e assinatura da formalização do projeto seguem somente após o aval do TAP e principalmente do orçamento.

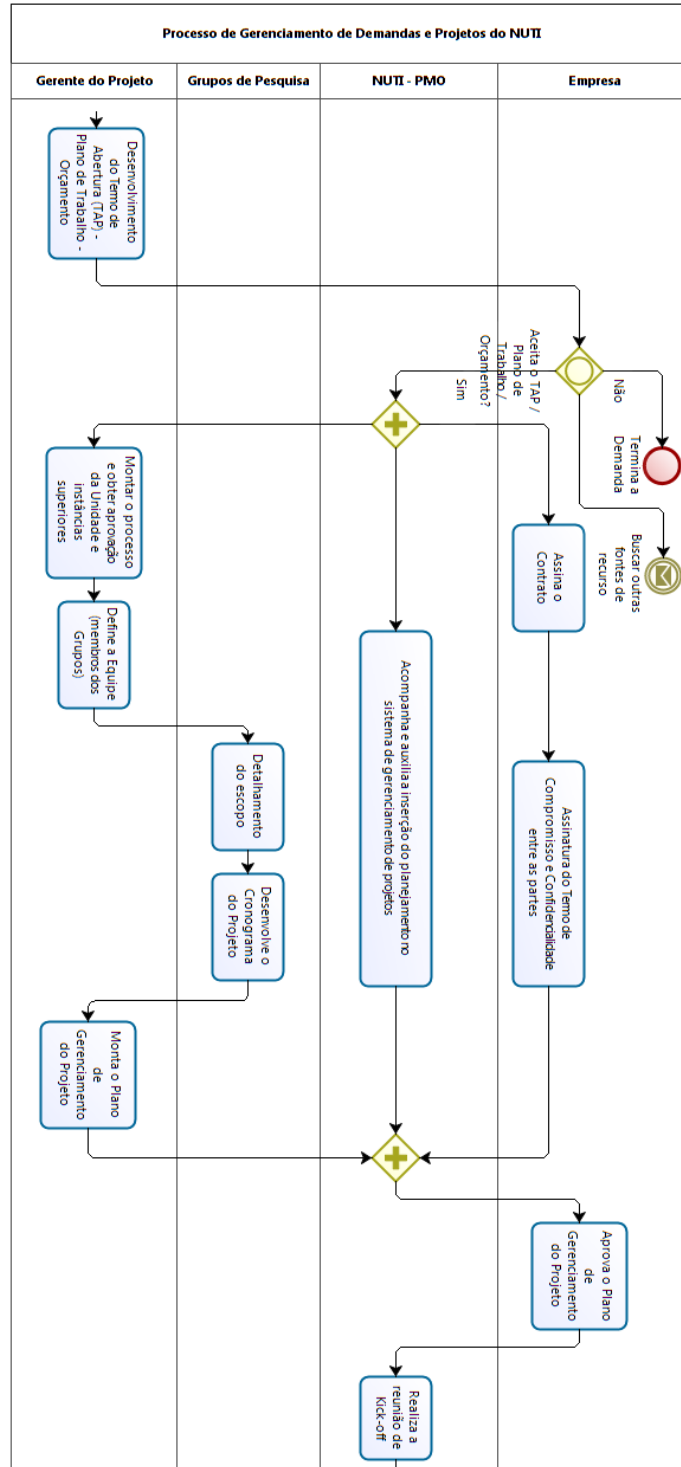


Figura 1- Fase de planejamento do projeto.

Ao final do planejamento, acontece a reunião de *kick-off*, é a partir deste momento que se dá início a execução do projeto.

O NUTI (PMO-IC) auxilia a comunicação entre os *stakeholders* e mantém sempre informado sobre o status do cronograma e demais áreas de conhecimento influenciadas pelo projeto.

### 3.3 Execução, Monitoramento e Encerramento

As fases de monitoração e de execução ocorrem paralelamente, finalizadas pelos diagnósticos trazidos na etapa de encerramento. Com o início da execução dos planos estipulados para o projeto por parte do grupo de pesquisa desde a reunião de *kick-off*, o gerente de projeto fica no encargo da gerencia do desenvolvimento e do grupo de pesquisa, nesta etapa o escritório de projetos deve disponibilizar os indicadores do projeto para o gerente do projeto para que seja possível o gerente criar a estratégia de andamento para com o projeto, com a estratégia e a rotina do projeto definida o grupo de pesquisa fica então responsável de desenvolver o projeto de acordo com o proposto. Nesse momento também, o grupo de pesquisa é auxiliado pela empresa solicitante que ditará as regras de negócios e fará vistorias nos resultados parciais durante o andamento do projeto, e fazendo a aprovação da mesma para prosseguir de acordo com o objetivo estipulado na etapa de desenvolvimento. Nesse mesmo tempo o PMO acompanha o desenvolvimento dos trabalhos utilizando os indicadores fornecidos pelo gerente do projeto, tendo isso como base para a coordenação a preparação dos relatórios ou produto desenvolvido pelo projeto ao final do processo de execução.

Com o encerramento pratico do projeto entramos na etapa de encerramento, onde temos uma reunião da equipe para discussão das lições aprendidas e a desmobilização da equipe. Com isso o projeto é encerrado.

## 4. Conclusão

A metodologia proposta tem a possibilidade de trazer uma grande melhoria ao ser utilizada em conjuntos com os projetos de pesquisas, podendo ser um importante fator de controle e organização sobre vários grupos de pesquisa que surgem no ambiente educacional.

Sua implementação não é simples da forma que aparenta ser, um estudo aprofundado sobre a matéria deve ser visto de antemão entre os envolvidos durante a criação do escritório de projetos para que haja uma completa formação e capacidade no momento de gestão dos projetos. Aplicar essa metodologia com uma equipe de PMO irá auxiliar aos envolvidos na pesquisa focar nas atividades relacionadas aos assuntos da pesquisa, utilizando menos esforço no gerenciamento do projeto, estes com responsabilidade da equipe do PMO.

### Referências bibliográficas

CLELAND, David I et al. Gerência de projetos. Reichmann & Affonso Editores, 2002.

CRAWFORD, J. Kent - The Strategic Project Office: A Guide to Improving Organizational Performance, 2001.

DINSMORE, Paul C. Winning in Business with Enterprise Project Management. New York: AMACON, 1999.

DINSMORE, Paul C. Human Factors in Project Management, Revised Edition. New York: AMACON, 1990.

KERZNER, H. Gestão de projetos: As melhores práticas. Porto Alegre: Bookman, 2006.

UNIVERSIDADE FEDERAL DO ESPIRITO SANTO - Grupos de Pesquisa - CNPq/UFES. Disponível em:< <http://www.prppg.ufes.br/grupos-de-pesquisa-cnpqufes>>. Acesso em: 27 de ago. 2014.

# Automação do Processo de Implantação de Software

Alicenaira L. C. Carneiro de Almeida<sup>1</sup>, Evaldo de Oliveira da Silva<sup>1</sup>,  
Daves Marcio Silva Martins<sup>1</sup>

<sup>1</sup>Centro de Ensino Superior de Juiz de Fora (CESJF)  
Caixa Postal 36.016-000 – Juiz de Fora – MG – Brazil

{alicenaira, evaldo.oliveira, davesmartins}@gmail.com

**Abstract** *In many environments of development the processes performed, between an issue for maintenance and deployment of the same production, are performed manually. The manual process of delivering software leads to production of software with longer cycles to generate a new version of the application. This article presents a methodology for automating the deployment of software applications based on the concepts of continuous delivery process.*

**Resumo:** *Em muitos ambientes de desenvolvimento os processos realizados, entre a solicitação de uma manutenção e a implantação da mesma em produção, são executados manualmente. O processo manual de entrega de software leva a produção de software com ciclos mais longos para geração de uma nova versão da aplicação. Este artigo apresenta uma proposta metodológica para automação do processo de implantação de aplicações de software com base nos conceitos de entrega contínua.*

## 1 Introdução

Geralmente equipes de software executam processos manuais de entrega de software, aumentando os riscos de erros durante a implantação em produção. Em muitos ambientes de desenvolvimento os processos não automatizados ou semi-automatizados entre o desenvolvimento e a implantação de uma solicitação, faz com que a duração do atendimento de demandas seja longo. Pode ocorrer a falta de controle sobre algumas atividades do ciclo de vida da aplicação, ou até mesmo o esquecimento da execução de algumas delas.

De acordo com Humble e Farley (2014) a entrega contínua de software envolve um conjunto de atividades de manutenção e implantação de software, possibilitando a confiabilidade e a previsibilidade dos processos de desenvolvimento. A entrega contínua também permite a repetição destes processos de forma automatizada, interligando as atividades de desenvolvimento, desde o versionamento do código até a implantação em ambiente de produção.

A automação da entrega de software pode ser obtida através de ferramentas e *scripts* que integram todas as atividades da entrega de software. Alguns processos de desenvolvimento podem ser integrados para atender a entrega contínua, como o Controle de Mudanças, o Controle de Versões, *build* de aplicações, testes automatizados, e a Gerência de Infraestrutura, necessitando do envolvimento de toda a equipe do escritório de software (SATO, 2014).

A ideia de automatizar as entregas de software serve para reduzir o tempo de implantação, permitindo que as equipes de software aumentem a frequência das entregas.

Porém, a criação de um processo de entrega é complexo e exige um conhecimento multidisciplinar que envolve áreas de Engenharia e Arquitetura de Software, Sistema Operacional e Infraestrutura de Tecnologia da Informação.

Desta forma, este artigo apresenta uma proposta metodológica para automação do processo de implantação de software com base nos processos de entrega contínua. Finalmente, descreve um conjunto de ferramentas utilizadas na indústria de software que podem fazer parte do processo para automação da implantação de software com base na proposta abordada.

O artigo está organizado com seções a seguir. A seção 2 apresenta os processos e ferramentas que apoiam a entrega contínua. A seção 3 a metodologia proposta neste trabalho. A seção 4 descreve como a metodologia pode ser utilizada em ambiente de desenvolvimento de software. A seção 5 apresenta as considerações finais e trabalhos futuros.

## **2 Processos e Ferramentas que apoiam a Entrega Contínua**

As próximas seções abordam os principais conceitos que fundamentam a proposta deste artigo.

### **2.1. Gerência de Configuração de Software**

A Gerência de Configuração de Software (GCS) é definida como uma área da Engenharia de Software responsável pelo controle da evolução de sistemas complexos. A GCS fornece técnicas, métodos e ferramentas que permitem a rastreabilidade e o monitoramento das manutenções de software (DART, 1991)

De acordo com Estublier (2000) os principais sistemas da GCS são o Controle de Mudanças, o Controle de Versões e a Gerência de Construção (ou *build*). Cada ferramenta serve para automatizar processos de desenvolvimento e integração de uma aplicação. Abaixo segue o objetivo de cada ferramenta de GCS e o processo ao qual está inserida:

- Sistemas de Controle de Mudanças: permitem acompanhar, auditar e gerenciar todas as solicitações de manutenção de software durante o ciclo de vida da aplicação.
- Sistemas de Controle de Versões: tem o objetivo de controlar a evolução dos itens de configuração do software com base em modelos de versionamento. Algumas ferramentas possuem arquitetura centralizada ou distribuída, e ambas permitem o desenvolvimento paralelo de software no atendimento das solicitações.
- Sistema de Construção: são conhecidas como ferramentas de *build*. Correspondem ao processo de compilação, teste e empacotamento da aplicação, envolvendo também o gerenciamento das dependências entre módulos e componentes do software

A GCS também contempla o processo de integração contínua. O objetivo é garantir que cada atualização do código fonte não gere regressões e/ou falhas em um aplicativo em desenvolvimento (FOWLER e FOEMMEL, 2006).

## 2.2 Teste de Software

De acordo com Humble e Farley (2014), a atividade de teste de software deve garantir a qualidade a cada mudança realizada na aplicação. Deste modo, vários tipos de testes devem ser contemplados no processo de entrega contínua.

A execução sistemática de testes tem o objetivo específico de encontrar e remover o maior número possível de erros, evitando assim que eles sejam encontrados pelo cliente (PRESSMAN, 2002).

Vários tipos de testes podem ser automatizados visando contribuir para o processo de entrega contínua, principalmente os testes de unidade e testes funcionais. Os testes automatizados são os responsáveis por garantir que quaisquer problemas que comprometam o cumprimento dos requisitos sejam identificados enquanto o custo para corrigi-los ainda é baixo, por isso devem ser bem definidos.

## 2.3 Gerência de Infraestrutura

Muitas aplicações de software possuem arquiteturas que dependem da configuração do hardware dos servidores, da infraestrutura de rede e do sistema operacional. É por meio deste ambiente que é possível manter vários serviços que permitem a execução das aplicações (HAMBLE; FARLEY, 2014).

A automação do processo de entrega contínua depende da configuração da infraestrutura, e da especificação desejada entre os ambientes de homologação e produção. A configuração destes ambientes deve ser equivalente para que o processo automatizado da entrega contínua seja realizado com sucesso (SATO, 2014).

## 2.4 O Pipeline de Implantação

O *pipeline* de implantação descreve um conjunto de *scripts* que automatiza o processo de compilação, testes e implantação do software. Permite que o processo de entrega seja realizado para cada versionamento de código-fonte, além do monitoramento e da medição do ciclo de desenvolvimento da aplicação (HUMBLE; FARLEY, 2014).

Com isso é possível que sejam identificadas melhorias na automação de testes ou na construção de uma nova versão do sistema. Quando uma versão passa por todo o *pipeline* com sucesso, significa que ela foi submetida a testes automatizados em ambiente similares aos de produção (HUMBLE; FARLEY, 2014).

## 3 Automação da Entrega Contínua de Software

Com o propósito de definir e documentar a entrega contínua de software, foi elaborada uma metodologia que descreve os passos para a automação do processo de desenvolvimento de software. A proposta está fundamentada nos seguintes passos:

- Definição do Sistema para Controle de Mudanças (SCM).
- Definição do Sistema de Controle de Versão, preferencialmente integrado ao SCM.
- Identificação dos componentes da arquitetura da aplicação e suas dependências.
- Configuração do servidor para a integração entre os módulos dos projetos.
- Definição dos procedimentos para analisar a qualidade do código.
- Determinação dos tipos de testes necessários.

- Configuração e definição de um gerenciador de infraestrutura.
- Criação do *pipeline* de implantação.

A próxima seção descreve as ferramentas e os processos que podem ser automatizados a fim de estabelecer a entrega contínua de software com base na proposta apresentada

### 3.1 Definição de ferramentas para automação da Entrega Contínua

A metodologia descreve a necessidade de escolher ferramentas de controle de mudanças que estejam integradas com o versionamento de software. A RedMine (BERTINO, 2012) e o Trac (MURPHY, 2007) permitem o monitoramento de mudanças integradas com o SubVersion ou SVN (BERTINO, 2012). O SVN é um sistema de controle de versão que atende ao segundo passo da proposta apresentada, permitindo a rastreabilidade entre o versionamento do código com a solicitação registrada no controle de mudanças.

O próximo passo é a definição da ferramenta para *build*. Este tipo de ferramenta permite o gerenciamento e compartilhamento da biblioteca de componentes do projeto. Algumas tecnologias de desenvolvimento permitem o *build* automatizado como, por exemplo, o Maven, o Gradle e o Ant (MUSCHKO, 2014).

Após a definição da ferramenta de construção é necessário a configuração do servidor para a integração dos módulos do projeto. O Jenkins (SMART, 2011) é uma ferramenta que permite esta integração, e possui interface com mais de 600 *plugins* para customização da compilação, análise, testes e implantação das aplicações. O Jenkins é responsável por interligar as ferramentas citadas anteriormente para a entrega de software por meio da criação de *scripts* que são executados no servidor.

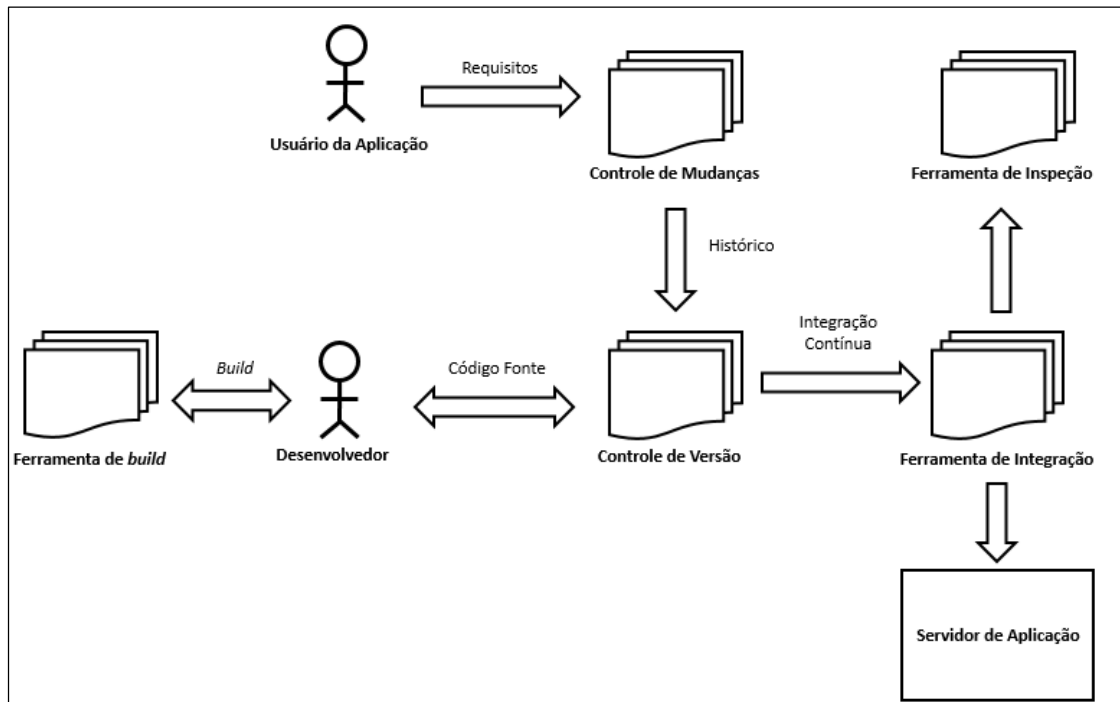
A ferramenta Sonar (ARAPIDIS, 2012) fornece *plugins* para o controle e medição da qualidade do código, e pode ser customizada de forma integrada ao Jenkins. As métricas são customizadas de acordo com a necessidade de cada projeto. É possível, por exemplo, definir um limite mínimo aceitável de incidentes por projeto.

A próxima etapa para definição do processo de entrega contínua, é a definição dos tipos de testes. Algumas tecnologias podem ser utilizadas para automatização de testes como o JUnit (RAINSBERGER, 2005) para a criação de testes unitários e a ferramenta Selenium (HOLMES, 2006) para a criação de testes funcionais.

Para o gerenciamento da infraestrutura deve-se utilizar uma ferramenta que permita reduzir o tempo de configuração entre os ambientes homologação e produção. O Puppet (LOOPE, 2011) é uma ferramenta que centraliza as configurações em um único ponto, permitindo que sejam distribuídas para outros nós dentro de uma rede. Utiliza linguagem declarativa para a configuração e execução de pacotes de programas em sistemas operacionais, entre eles Linux, Windows e Solaris.

A última etapa da proposta apresentada, é a criação do *pipeline* de implantação no servidor de integração dos módulos do projeto. É importante ressaltar que não existe uma solução única para a criação de um *pipeline*, as configurações podem variar de acordo com a complexidade e arquitetura de cada projeto. Por meio da ferramenta Jenkins é possível criar o *pipeline* de forma automática interligando os *scripts* de execução do processo de entrega contínua, podendo ser disparados de forma manual ou automática.

A Figura 1 representa o processo de entrega contínua com a utilização das ferramentas descritas para o processo de automação.



**Figura 1. Representação do processo de Entrega Contínua**

O processo representado na Figura 1 apresenta o ciclo de modificação da aplicação desde da solicitação de mudança realizada pelo usuário até a integração da solução de software e implantação no servidor de aplicação. Com isso, a cada versionamento do código é possível ter uma nova entrega de forma automatizada, obedecendo os requisitos registrados pelo usuário. É importante ressaltar que todo processo pode ser automatizado, ou parte dele. Por exemplo, a etapa de implantação no servidor de aplicação e a inspeção do código podem ser manuais.

#### 4 Considerações Finais

A metodologia apresentada possibilita a automação dos principais processos da entrega de software e pode ser utilizada por qualquer equipe que pretenda implantar a entrega contínua. A escolha das ferramentas pode variar de acordo com o escopo e a necessidade de cada projeto, cabe a cada equipe avaliar quais são as melhores soluções.

É importante ressaltar que existem muitas opções de ferramentas para automatizar uma atividade do processo de entrega contínua. Com a abordagem proposta é possível substituir uma ferramenta sem ter que o processo de entrega automatizada.

Como trabalhos futuros, pretende-se criar mecanismos automatizar o processo de implantação em outros ambientes, como, por exemplo, aplicações que utilizam serviços na nuvem. Além disso, é necessário o envolvimento de toda a equipe de software visando mudanças culturais para ajudar na melhoria contínua dos processos com o desenvolvimento colaborativo utilizando a entrega contínua.

## Referências

- ARAPIDIS, Charalampos. Sonar Code Quality Testing Essentials. Packt Publishing Ltd, 2012.
- BERTINO, Nic. Modern version control: creating an efficient development ecosystem. In: Proceedings of the 40th annual ACM SIGUCCS conference. ACM, 2012. p. 219-222.
- DART, Susan. Concepts in configuration management systems, Proceedings of the 3rd international workshop on Software configuration management, p.1-18, June 12-14, 1991.
- ESTUBLIER, J. Software Configuration Management: a Roadmap. International Conference on Software Engineering (ICSE), The Future of Software Engineering. Limerick, Ireland. June, 2000. 279-289 p.
- FOWLER, Martin; FOEMMEL, Matthew. Continuous integration. Thought-Works [http://www.thoughtworks.com/Continuous Integration. pdf](http://www.thoughtworks.com/Continuous%20Integration.pdf), 2006. Acesso em 12 de Ago de 2014.
- LOOPE, James. Managing Infrastructure with Puppet. " O'Reilly Media, Inc.", 2011
- HOLMES, Antawan; KELLOGG, Marc. Automating functional tests using selenium. In: Agile Conference, 2006. IEEE, 2006. p. 6 pp.-275
- HUMBLE, Jez; FARLEY, David. Entrega contínua: como entregar software de forma rápida e confiável. Porto Alegre: Bookman, 2013.
- MURPHY, David J. Managing Software Development with Trac and Subversion. Packt Publishing Ltd, 2007.
- MUSCHKO, Benjamin. Gradle in Action. Manning, 2014.
- PRESSMAN, Roger S. Engenharia de Software. 5. Ed. São Paulo: Ed. McGraw-Hill, 2002.
- RAINSBERGER, Joe B. JUnit recipes: practical methods for programmer testing. Manning, 2005.
- SATO, Danilo. DevOps na prática: entrega de software confiável e automatizada. São Paulo: Caso do Código, 2014.
- SMART, John. Jenkins: the definitive guide. " O'Reilly Media, Inc.", 2011.

**Caderno de**  
**Resumos Expandidos**



## A Implementação de monitoramento centralizado em servidores do Instituto de Computação da UFMT

César Eduardo Guarienti<sup>1</sup>, Nielsen Simões<sup>2</sup>

1, 2 – Instituto de Computação – Universidade Federal de Mato Grosso (UFMT) - Cuiabá - MT

{cesar.guarienti, nielsen}@ic.ufmt.br

**Abstract.** *This article presents a case study of centralized resources for servers was successfully implemented at the Institute of Computing (IC) of the Federal University of Mato Grosso (UFMT), Cuiabá-MT campus monitoring. To provide multiple services on a computer network and multiple servers are created with the growth of the number of machines controlling these resources is increasingly difficult. Without the power of the information management of these resources may be inaccurate and lead to unnecessary expenses. This study demonstrates that it is possible to centrally monitor multiple servers, generate graphs with different types of information and provide this data on the Internet.*

**Resumo.** *Este artigo apresenta um estudo de caso de monitoramento centralizado de recursos para servidores que foi implementado com sucesso no Instituto de Computação (IC) da Universidade Federal de Mato Grosso (UFMT), campus Cuiabá-MT. Para fornecer diversos serviços em uma rede de computadores vários servidores são criados e com o crescimento do número de máquinas o controle desses recursos é cada vez mais difícil. Sem o poder dessas informações o gerenciamento desses recursos pode ser impreciso e levar a gastos desnecessários. Esse estudo demonstra que é possível monitorar de forma centralizada vários servidores, gerar gráficos com diversos tipos de informações e disponibilizar esses dados na Internet.*

### 1. Introdução

O número de serviços oferecidos em uma rede de computadores é cada vez maior. A quantidade de dispositivos que se integram com as redes é crescente. Usuários da tecnologia da informação têm necessidades cada vez mais sofisticadas e anseiam por respostas imediatas. A resposta dessas demandas, dessa nova realidade exige uma organização maior das informações a respeito dos recursos disponíveis. De acordo com STALLINGS (1999) o custo pra realizar essa atividade de gerenciamento dos recursos muitas vezes pode chegar até a 15% dos recursos investidos na área de tecnologia da informação.

A administração do parque tecnológico requer dados precisos. Para a tomada de decisão sobre a implantação de um serviço novo ou uma nova tecnologia é necessário verificar se estão disponíveis os requisitos mínimos exigidos em cada caso. Um

exemplo é a tecnologia de virtualização que oferece vários benefícios como economia de energia e a possibilidade de fazer balanceamento de carga com menor esforço, além disso, propicia que se tenha vários servidores em um só *host*, mas por outro lado torna difícil organizar as informações sobre servidores que estão em lugares diferentes. Outro problema é que o uso de virtualizadores de fabricantes diversos é cada vez mais comum. Nesse cenário com fabricantes diferentes e máquinas localizadas em pontos diversos é importante obter informações rápidas e fáceis pra poder realizar ajustes no ambiente. Através de gráficos e relatórios podemos fazer ajustes finos, realocar disco e memória conforme a demanda, por exemplo. Segundo ZARPELÃO (2004) o monitoramento de ativos de rede contribui para o funcionamento e confiabilidade das redes de computadores.

Para monitorar e analisar a performance dos recursos de tecnologia da informação disponíveis no Instituto de Computação da Universidade Federal de Mato Grosso, campus Cuiabá, foi implantado o *munin* (<http://munin-monitoring.org>), ferramenta gratuita de código livre.

Esse artigo apresenta um estudo de caso de monitoramento centralizado de servidores usando gráficos e relatórios que fornecem várias informações sobre diversos componentes, redes e serviços, como *apache* e *postfix*.

## 2. Metodologia

O cenário é formado por *hosts* que utilizam virtualizadores de diferentes fornecedores, localizados em diversos pontos. Nesses equipamentos estão máquinas virtuais de código aberto e fechado. Para que essas máquinas virtuais hospedadas pudessem comunicar com o *munin server* foi necessário instalar os componentes do pacote cliente, além disso, cadastrar no servidor todos os recursos que deveriam ser monitorados.

Para que as informações geradas pudessem ser visualizadas de qualquer lugar o servidor foi colocado no *proxy* reverso do IC. Assim, foi possível consultar os dados pelo endereço: <http://monitor.ic.ufmt.br> através da internet. Na Figura 1 pode ser visualizado a tela inicial do monitoramento com os clientes que foram cadastrados, é possível observar algumas opções dos clientes cadastrados como: disco, rede, processamento e sistema. Ao selecionar essas opções uma série de gráficos são apresentados com diversos tipos de informações diferentes.



Figura 1: Clientes Cadastrados

### 3. Resultados

A solução apresentada foi capaz de coletar e apresentar as informações desejadas. Através disso foi possível monitorar recursos através de uma série de gráficos detalhados, mesmo pela internet, de vários servidores localizados em pontos diferentes do IC-UFMT.

A partir das informações coletadas foi possível fazer um remanejamento de máquinas virtuais que tinham consumo maior de disco e memória, o que resultou na desativação de *hosts* antigos que possuíam configuração defasada.

Na Figura 2 é possível verificar a porcentagem de uso do disco de um servidor virtual que foi remanejado por conta do alto volume de dados.

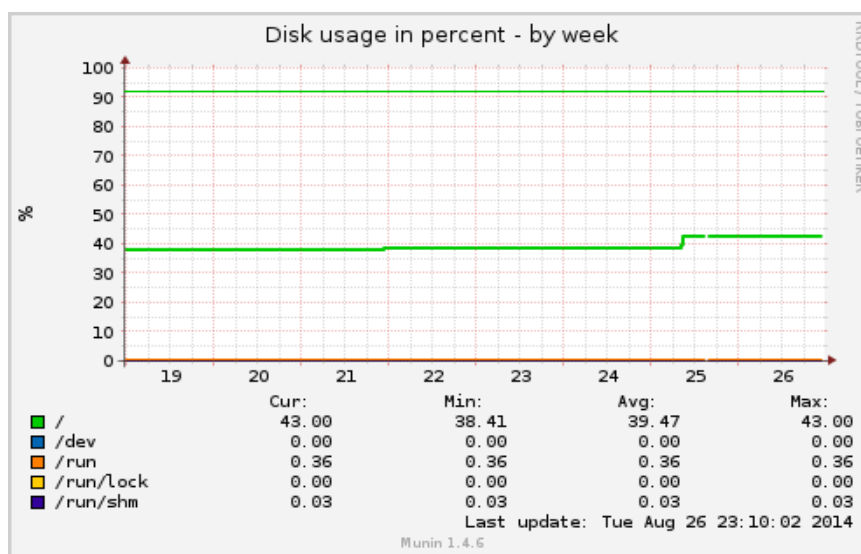


Figura 2: Porcentagem de uso de disco

### 4. Considerações Finais

O gerenciamento de recursos de tecnologia da informação é fundamental pra dar uma resposta positiva aos anseios cada vez mais complexos dos usuários das redes de computadores. Para implantar um serviço novo muitas vezes é necessário conhecer o cenário atual antes de investir em uma nova infraestrutura. O monitoramento de recursos é uma alternativa viável de se obter diversas informações sobre o *hardware* disponível e a partir disso oferecer elementos para a tomada de decisão sobre investimentos em tecnologia da informação, por exemplo, investir mais em armazenamento de dados.

### 5. Referências

- STALLINGS, W. SNMP, SNMP v2, SNMP v3 and RMON 1 and 2, 3ª edição, Addison - Wesley, 1999.
- ZARPELÃO, B. B. Detecção de Anomalias e Geração de Alarmes em Redes de Computadores, Trabalho de Conclusão de Curso de Ciência da Computação, Universidade Estadual de Londrina, 2004.

## A utilização de dicionário de sinônimos na análise automática de patentes

Bruna Cruz Carlota<sup>1</sup>, Fernando Muniz da Cruz<sup>1</sup>, Andreia Gentil Bonfante<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)

brunacarlota@gmail.com, fernandodnto@gmail.com, andreia.bonfante@ic.ufmt.br

**Abstract.** *This paper presents a initial research using thesaurus and text mining techniques in order to improve performance of information retrieval and automatic categorization. This module will be part of the Periscope, an open source patent analysis tool.*

**Resumo.** *Este artigo apresenta avanços iniciais do estudo da viabilidade da utilização de dicionário de sinônimos junto com técnicas de processamento de texto com o propósito de melhorar a eficiência da recuperação de informações e categorização automática. Essa funcionalidade fará parte do Periscope, uma ferramenta baseada em software livre que auxilia na análise de patentes.*

### 1. Introdução

Com o intuito de competir por mercado ou gerar novas tecnologias, os pesquisadores necessitam a todo momento acessar o conteúdo de documentos de patentes, seja para observar uma tecnologia existente ou para fundamentar suas invenções. O crescimento constante dessas bases, com quantidades cada vez maiores de documentos, faz com que haja a necessidade da ajuda de ferramentas capazes de produzir análises, como tendências tecnológicas, identificação de áreas promissoras de pesquisa, extração de informação para detecção de violações, etc. Dentre as tarefas de análise estão a visualização, as análises de citação, e a aplicação de técnicas de mineração [Abbas et al. 2014]. As técnicas de mineração de textos produzem uma análise *full-text* em conteúdos não estruturados das patentes, como título, resumo, produtos relacionados, descrição, etc., e trabalham seguindo algumas etapas que são: pré-processamento dos documentos (criação da coleção, etc.); indexação; agrupamento por tópicos (seleção de termos, categorização, etc.); e mapeamento (de tendências, agregação, etc.) [Tseng et al. 2007].

A indexação tem como objetivo extrair os termos que identificam os documentos e armazená-los de uma forma que sejam recuperados facilmente. Para que haja a indexação dos textos, é necessário também que na fase de pré-processamento os documentos passem por um processo de tokenização, que consiste na retirada dos espaços e caracteres especiais do texto, tendo como produto apenas as palavras. No pré-processamento pode haver também a remoção de *stopwords*, palavras consideradas irrelevantes para o conjunto de resultados que se deseja obter. São palavras como: *para, com, a, o, as, os, mas*, etc. O produto deste pré-processamento é o conjunto de palavras já prontos para serem indexados [Mccandless et al. 2010].

No entanto, considerando cenários de aplicação como busca por similaridade ou categorização/classificação automática, os resultados não são tão satisfatórios, pois não

são capazes de lidar com questões básicas envolvendo sinonímia, ou seja, vários textos podem estar falando do mesmo assunto, mas as pesquisas apenas referenciarão os documentos que contêm o termo buscado.

Neste artigo, são apresentados os avanços iniciais para o estudo da viabilidade da incorporação de um dicionário de sinônimos ao processamento automático desses documentos de patentes, no intuito de colaborar na categorização automática e na busca por similaridade. Tendo como foco o processamento de textos na língua portuguesa do Brasil, está sendo implementado um módulo que integra a base WordNet.Br a um sistema de indexação de textos de resumos de patentes. O módulo fará parte de uma aplicação, chamada Periscope, uma ferramenta de análise de patentes, que está sendo desenvolvida por um grupo de pesquisa do qual os autores fazem parte <sup>1</sup>, em parceria com a Academia do Instituto Nacional de Propriedade Industrial (INPI). Para a implementação, está sendo utilizado o Apache Lucene<sup>TM</sup>, uma biblioteca para indexação e busca *full-text* de documentos textuais, da Apache Software Foundation.

## 2. Referencial Teórico

Desenvolvido pela Universidade de Princeton, a Wordnet é uma eficiente base de dados léxicos que reúne grande quantidade de verbos, substantivos, adjetivos e advérbios e os associa a sinônimos, formando os conjuntos de sinônimos denominados *SynSets* (*Synonym Sets*). Tais conjuntos são relacionados de acordo com a semântica que apresentam, criando assim, uma rede de palavras conceitualmente relacionadas, em que palavras de uma classe (adjetivos, substantivos, advérbios e verbos) só se relacionam com palavras da mesma classe [Scarton and Aluísio 2009].

Inicialmente, essa ferramenta era exclusivamente disponibilizada para a língua inglesa, mas por ter se mostrado bastante eficiente, pesquisadores começaram a estudar formas de adaptar a WordNet para outros idiomas, dentre os quais o português brasileiro, em um projeto denominado WordNet.Br ([Dias-da silva and Ávila 2005], [Dias-da silva 2003]), contando atualmente com 19.882 *SynSets* [Oliveira and Roman 2013]. Esta base funcionava conectada à WordNet.Pr (oficial de Princeton<sup>2</sup>) e um dicionário bilíngue ao qual os termos buscados eram convertidos para seu significado em inglês e submetidos a uma busca na WordNet.Pr, que trazia todas as relações de hipernímia encontradas. Em seguida, a WordNet.Br convertia todos os termos encontrados novamente para o português e o usuário que solicitou a busca poderia visualizá-las.

No entanto, estas conversões de um idioma para outro podem prejudicar as buscas, pois estão sujeitas a perdas de informações. Uma das formas encontradas para resolver este problema foi criar uma API de acesso à WordNet local, tal como a API utilizada na WordNet.Pr, a API JAWS. A API Java para a WordNet.Br é a JWN-Br [Oliveira and Roman 2013]. Esta API faz com que um objeto implemente a interface *SynSet*, cuja função é fornecer os métodos de acesso a cada um de seus elementos. A WordNet.Br contém um arquivo que guarda informações necessárias para a construção de um *SynSet*. Estas informações são: a palavra principal, o número referente àquela palavra, a função sintática, um conjunto de sentidos, uma lista de sinônimos e uma lista

<sup>1</sup>Com o apoio da UFMT, do CNPq e da FAPEMAT

<sup>2</sup><http://wordnet.princeton.edu/>

de antônimos. Uma palavra pode ter mais de um sentido e aparecer em contextos diferentes. Esta ferramenta traz como resultado de pesquisa todos os contextos em que determinado termo pode estar inserido. Os principais resultados apresentados são os conjuntos de sinônimos, mas um conjunto de antônimos está associado a cada termo.

A integração da WordNet.Br à aplicação aqui descrita pode ser feita de duas maneiras: na indexação dos termos, e na busca. A indexação dos termos é teoricamente simples. Antes de ser indexado, ocorre uma busca desse termo no conjunto de *SynSets*, e o resultado é adicionado na indexação junto com o termo original, como se todas as palavras ocupassem apenas um espaço no índice. Então, quando a busca é realizada, as palavras que foram indexadas junto com a original serão buscadas igualmente nos documentos. Já quando utilizada na busca, a consulta é composta pela palavra original e seus sinônimos, recuperados da base.

### 3. Trabalhos a serem realizados

O protótipo que validará esta pesquisa está em fase inicial de desenvolvimento. Já foi feita a associação com a WordNet.Br, que se dá na indexação das palavras. Assim, cada palavra de cada documento é indexada juntamente com seus sinônimos, recuperados da base. Um módulo de recuperação por palavra também foi implementado, como primeira forma de analisar os resultados. Alguns problemas de codificação da base e dos textos ainda devem ser superados, uma vez que a indexação e a busca são altamente dependentes de particularidades como acentuação, e devem ser padronizados. Esforços estão sendo feitos no sentido de uniformizar as codificações. Após essa fase inicial de integração da WordNet.Br ao módulo de indexação, é necessário o aperfeiçoamento deste processo, com análises consistentes de sua viabilidade quando utilizados em uma aplicação real. A integração deste a um módulo de categorização automática, também já em fase de implementação, será o próximo passo.

### References

- Abbas, A., Zhang, L., and Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information*, 37(0):3 – 13.
- Dias-da silva, B. C. (2003). A construção da base da Wordnet.Br : conquistas e desafios. In *XXV Congresso da Sociedade Brasileira de Computação*, pages 2238–2247, São Leopoldo, RS, Brasil. SBC.
- Dias-da silva, B. C. and Ávila, M. C. (2005). Alternâncias sintáticas, estrutura de argumentos, synsets e a rede Wordnet.Br. *Estudos Linguísticos XXXIV*, 1(1):368–373.
- Mccandless, M., Hatcher, E., and Gospodnetic, O. (2010). *Lucene In action*. Manning Publications Co., second edition.
- Oliveira, V. M. and Roman, N. T. (2013). Jwn br: uma api java para a wordnet.br. In *Proceedings of the 9th Brazilian Symposium in Information and Human Language Technology*, pages 153–157, Fortaleza, CE, Brazil. Sociedade Brasileira de Computação.
- Scarton, C. E. and Aluísio, S. M. (2009). Herança automática das relações de hiperonímia para a WordNet.Br. Technical report, ICMC-USP, São Carlos, SP, Brasil.
- Tseng, Y.-H., Lin, C.-J., and Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information Processing and Management*, 43(5):1216 – 1247. Patent Processing.

# A utilização de medida de similaridade na classificação automática de documentos de patentes

Allan Ferreira Geraldo de Alencar<sup>1</sup>, Rafaely Fernanda Lino<sup>1</sup>,  
Andreia Gentil Bonfante<sup>1</sup>

<sup>1</sup>Instituto de Computação – Universidade Federal de Mato Grosso (UFMT)

allan.alencar.1995@hotmail.com, rafaelylino@gmail.com

**Abstract.** *This paper presents the initial progress in building a module for automatic classification of patent documents. This functionality will be part of the Periscope, an open source patent analysis tool.*

**Resumo.** *Este artigo apresenta os avanços iniciais na construção de um módulo de classificação automática de documentos de patentes. Essa funcionalidade fará parte do Periscope, uma ferramenta baseada em software livre que auxilia na análise de patentes.*

## 1. Introdução

As bases documentos de patentes são fontes necessárias de consulta para especialistas interessados em fundamentar suas invenções, analisar tendências de mercado, identificar pesquisas promissoras, etc. Técnicas automáticas de análise de patentes são utilizadas para que o processo de busca e recuperação de informação seja facilitado [Abbas et al. 2014]. Algumas delas utilizam similaridade, com o propósito de agrupar (processo chamado de *clustering*) ou classificar documentos automaticamente [Maia and Souza 2008]. Elas permitem que conteúdo semelhante seja recuperado e também auxiliam na organização dentro de um sistema de hierarquia.

Seu princípio de funcionamento parte da premissa de pré-processar os documentos disponíveis de tal forma que deles sejam extraídos descritores que os representam, para então indexá-los, facilitando uma recuperação posterior. Na maioria das aplicações, esses descritores são uma lista de palavras extraídas dos documentos. Uma das formas mais eficientes de indexação, utilizada neste trabalho, é a de índice invertido, que indica em qual documento cada palavra aparece.

Dentre os modelos clássicos de estruturação, está o modelo vetorial. Ele representa as palavras de cada documento e seus respectivos pesos em forma de vetores, e assim, permite que sejam localizados documentos similares. Os pesos são utilizados para calcular o grau de similaridade entre a consulta e o documento.

Este artigo apresenta avanços iniciais na construção de um módulo de classificação automática de documentos de patentes escritas na língua portuguesa brasileira. O módulo fará parte de uma aplicação, chamada Periscope, uma ferramenta de análise de patentes, que está sendo desenvolvida por um grupo de pesquisa do qual os autores fazem parte <sup>1</sup>,

<sup>1</sup>Com o apoio da UFMT, do CNPq e da FAPEMAT

em parceria com a Academia do Instituto Nacional de Propriedade Industrial (INPI). Para a implementação, está sendo utilizado o Apache Lucene™, uma biblioteca para indexação e busca *full-text* de documentos textuais, da Apache Software Foundation.

## 2. Referencial Teórico

Todos os conjuntos de documentos necessitam de uma representação, tal que o uso de processos de mineração de textos como a coleta, o pré-processamento, a indexação, a busca e a visualização, possam atuar. Dessa forma, é necessário encontrar métodos que sugiram organizações automáticas desses documentos.

A medida de similaridade é fundamental para essa organização, seja para adicionar novos documentos a uma determinada classe, ou até mesmo para retirá-los. Dentro do Sistema de Recuperação de Informação (SRI), especificamente nos modelos clássicos de indexação, destacam-se o booleano, o vetorial e o probabilístico. O modelo mais utilizado que permite encontrar a similaridade entre dois documentos é o modelo vetorial, que se define a partir de um conjunto de documentos que o forma. Todas as palavras do texto são extraídas e convertidas para um formato de fácil manipulação, sendo a cada palavra associado um peso, para então, formarem um vetor (termo por documento) [McCandless et al. 2010].

A estrutura do vetor é organizada da seguinte forma: cada linha representa uma palavra (termo), e cada coluna representa um documento. Para uma coleção de  $N$  documentos  $d = (d_1, d_2, \dots, d_n)$  que inclui  $M$  elementos diferentes  $t = (t_1, t_2, \dots, t_m)$ , cada documento  $d_i$  é um vetor  $d_i = (\alpha i_1, \alpha i_2, \dots, \alpha i_m)$ , no qual o valor de  $\alpha i_j$  é uma medida de ocorrência, ou frequência, do termo  $t_j$  no documento  $d_i$  (peso).

Existem alguns métodos para calcular essas medidas, tais como:

- Booleano, que utiliza apenas da indicação de pertencimento de um termo a determinado documento;
- Frequência do termo (*tf*), que calcula o número de vezes que um termo aparece em um documento;
- Frequência do termo inverso da frequência do documento *tf-idf*, que considera que um termo é relevante de acordo com sua frequência em cada documento do conjunto.

Para qualificar a semelhança entre a informação de dois documentos, algumas medidas de similaridade são utilizadas. As mais comuns são [Maia and Souza 2008]:

- Distância Euclidiana, baseada na diferença entre os pesos dos termos dos documentos. Nos casos em que a maioria dos termos tem peso com valor zero, a comparação faz com que ele classifique erroneamente dois documentos como similares;
- Distância de Manhattan, relacionada à distância euclidiana, correspondendo ao comprimento do caminho mais curto entre dois pontos no espaço.
- Similaridade do Cosseno, que permite fazer comparações do tipo zero-zero; usada especialmente para ambientes de alta dimensionalidade e esparsos.

Neste trabalho, foi adotada a medida do cosseno. Ela consiste em vetorizar as palavras dos documentos e calcular o cosseno do ângulo que formam entre si. Quanto menor o ângulo, mais próximo de 1 é o cosseno e, conseqüentemente, mais similares são os documentos [Ingersoll et al. 2013].

### 3. Trabalhos a serem realizados

A implementação atual do módulo de classificação utiliza como base a hierarquia já pré-definida das classes de patentes. Para a vetorização, é utilizada a indexação de *term-vectors* do Lucene.

O pré-processamento realizado nos termos do documento aplica filtros de tokenização, transformação em minúsculas, remoção de acentuação e de *stopwords*<sup>2</sup>. Os tokens restantes são então vetorizados e associados a sua classe específica. Ao receber um documento consulta, o mesmo processo de pré-processamento é feito nele, e então é computado o cosseno do ângulo entre ele e os vetores que mapeiam cada classe.

Dentre os trabalhos a serem realizados estão o tratamento dos documentos, resolvendo problemas de codificação, trazidos na importação dos repositórios, testes e avaliações consistentes e representativas, utilizando o documento inteiro como consulta e a seleção dos termos mais representativos para a composição dos vetores.

### References

- Abbas, A., Zhang, L., and Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information*, 37(0):3 – 13.
- Ingersoll, G. S., Morton, T. S., and Farris, A. L. (2013). *Taming Text*. Manning Shelter Island, New York.
- Maia, L. C. G. and Souza, R. R. (2008). Medidas de similaridade em documentos eletrônicos 1. *IX ENANCIB: Diversidade Cultural e Políticas de Informação*, 1(1):1–15.
- Mccandless, M., Hatcher, E., and Gospodnetic, O. (2010). *Lucene In action*. Manning Publications Co., second edition.

---

<sup>2</sup>Palavras não discriminantes do documento.

## Legado Digital: novos paradigmas culturais

Rennan de Figueiredo, Lucas Sartorello, Aline P. Domingos de Souza<sup>1</sup>

<sup>1</sup>UNIRONDON– Centro Universitário Candido Rondon  
Avenida Beira Rio, 3001. Jardim Europa. 78.015-480 – Brazil

[rennan@agendaassessoria.com.br](mailto:rennan@agendaassessoria.com.br), [lucassartorello@gmail.com](mailto:lucassartorello@gmail.com)  
[aline.souza@pitagoras.com.br](mailto:aline.souza@pitagoras.com.br)

***Abstract.** The paper reflects the digital legacy and its various forms of representation in the virtual scenario, civil rights and the information that will be relevant. The choice of digital legacy as an object of research occurred because it is a conflicting relationship between the real and the virtual. In conclusion, will be presented, a different look at the cultural changes that contemporary society is experiencing specifically in relation to digital culture.*

***Resumo.** O trabalho reflete o legado digital e suas várias formas de representação no cenário virtual, os direitos civis e as informações que deverão ser preservadas. A escolha da herança digital como objeto de investigação ocorreu por se tratar de uma relação conflitante entre o real e o virtual. Como conclusão, será apresentado um olhar diferente sobre as mudanças culturais que a sociedade contemporânea esta vivenciando especificamente no que se refere à cultura digital.*

### 1. Introdução

No mundo contemporâneo, as novas tecnologias vêm influenciando a vida social, cultural e econômica. Como sugerido por André Lemos (2003) uma reestruturação através das novas ferramentas de comunicação toma conta da comunicação das mídias e das relações sociais. As redes sociais modificaram a maneira de como nos relacionamos ultrapassando, assim, os limites entre o real e o virtual.

A partir do momento que a interação homem-máquina tornou-se onipresente, através da computação ubíqua, houve uma modificação dos comportamentos, valores e costumes, gerando o que Lévy (1999) denominou de cibercultura.

Os ritos de enlutamento também passaram pelas modificações que o mundo virtual trouxe, o que enfatiza este novo cenário é a relação tempo e espaço, ou melhor, o ciberespaço, que segundo Willian Gibson tem como significado ser um espaço não físico ou territorial, que se compõe de um conjunto de redes de computadores através das quais todas as informações circulam, portanto o ciberespaço veio para se apresentar como um novo horizonte de expressão da sociedade.

Para Silvestre e Aguilera (2008), essa nova socialização ocorre porque a internet é uma expansão do mundo físico, e questões referentes à vida e morte migraram para o mundo virtual em busca de novos significados e de novas maneiras de expressão, em um mundo de constantes transformações.

### 2. Proposta e Resultados

O que fazer com o legado digital após a morte? Devido a grande produção de conteúdo, informações e dados com potencial de armazenamento o legado digital vem

despertando discussões. Os resultados dessa mudança de paradigmas já são vivenciados. O trabalho visa enfatizar reflexões que são muito recentes dentro do ciberespaço e assim compreender as novas formas de enlutamento que ocorreram e o seu impacto na sociedade contemporânea.

Segundo o pesquisador Cristiano Maciel, várias abordagens devem ser consideradas, dentre elas destacamos: a busca por soluções para a inserção do aspecto volitivo na produção destas redes sociais, desta decisão por parte do próprio usuário, não deixando esse legado para o herdeiro, como acontece hoje de acordo com os termos de uso das redes sociais (BORGES, L. C. L. F. ; FILGUEIRAS, L. V. L. ; MACIEL, C, 2011); investigaram se a crença religiosa do usuário influencia na sua decisão quanto ao legado digital, se quer manter o perfil ativo, se quer guardar os dados e excluir o perfil (MACIEL, Cristiano; PEREIRA, V.C., 2012) e o conceito de interação póstuma e como os jovens da geração internet lidam com o assunto. (PEREIRA, V.C. ; Maciel, Cristiano, 2012).

Atualmente, no Brasil não há leis em vigor que tratam o tema herança digital, apenas projetos de lei, como a “PL N.º 4.847, de 2012” elaborada pelo Deputado Marçal Filho, que tem como objetivo a elaboração das normas sobre herança digital no país e a PL n.º. 4099/2012 elaborada pelo Jorge de Mello que busca a Garantia aos herdeiros a transmissão de todos os conteúdos de contas e arquivos digitais.

No Brasil, o conceito de herança digital ainda é pouco utilizado. Mas é preciso uma legislação apropriada para que as pessoas ao morrerem possam ter seus direitos resguardados a começar pela simples decisão de a quem deixar a senha de suas contas virtuais e também o seu legado digital. Quando não há nada determinado em testamento, o Código Civil prioriza familiares da pessoa que morreu para definir herdeiros. Desta forma, os presentes Projetos de Leis pretendem assegurar o direito dos familiares em gerir o legado digital daqueles que já se foram.

Mais quais informações são relevantes para serem preservadas? Como e de quais maneiras deverão ser filtrados estas informações? Deverá ser criado um “Testamento Virtual”? No qual pode ser descrito e programado o que será feito com seus bens virtuais.

A partir da ideia que existe na Internet uma busca intensa pela conexão social, é prudente observar o quanto as relações online mantêm em sua essência valores que fundam os papéis e as relações sociais estabelecidas nos ambientes fora da rede. Assim, ao considerar as aproximações entre ambas, percebe-se que no ciberespaço, tal como nos outros ambientes socioculturais, são mantidas as teatralizações do cotidiano, os conflitos e as contradições da existência do sujeito perante o coletivo (LEMOS, 2003).

Atualmente, já existem sites que oferecem o serviço de fazer o seu testamento virtual no qual se pode escolher qual será o destino de todo o seu conteúdo virtual tais como contas de e-mails, sites, armazenamentos virtuais, contas de bancos e muito mais, essas informações serão disponibilizadas para as pessoas em que você listou após um tempo determinado de inatividade. Serviços como esse parece ser uma coisa estranha, porém ele pode decidir o destino de tudo o que foi feito por você ao longo de sua vida.

### **3. Conclusão**

O trabalho observou as mudanças de paradigmas que a sociedade vem de forma gradual passando.

Nos dias atuais a tecnologia está presente em todos os momentos de nossa vida, expressando uma forma de comunicação e relacionamento. É evidente que as novas formas de vivenciar e se expressarem após o advento da internet estão sendo modificadas. Ainda é cedo para determinar o quanto estas manifestações iram interferir nas construções simbólicas sobre a morte evidenciadas na cibercultura.

Algumas questões não foram trabalhadas, porém possuem grande potencial para futuras pesquisas tais como investigaram se a crença religiosa do usuário influencia na sua decisão quanto ao legado digital.

## Referencias

BORGES, L. C. L. F. ; FILGUEIRAS, L. V. L. ; MACIEL, C. . Towards a participatory development technioque of assistive technology for mobility and speech impaired patients. Anais do X Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais & V Conferência Latino-Americana de Interação Humano-Computador.. Porto Alegre: SBC, 2011. p. 247-256.

LEMOS, André. Cibercultura: alguns pontos para compreender nossa época. In: LEMOS, André;CUNHA, Paulo (orgs). Olhares sobre a Cibercultura. Sulina, Porto Alegre, 2003.

LÉVY, P. Cibercultura. São Paulo: Ed. 34, 1999.

Maciel, Cristiano ; PEREIRA, V.C. . The influence of beliefs and death taboos in modeling the fate of digital legacy under the software developers' view. In: ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2012), 2012, Texas.

PEREIRA, V.C. ; Maciel, Cristiano . The user's and the designer's role and the aesthetic experience of generative literature. Manaus-AM. Brazilian Symposium on Human Factors in Computing System. Porto Alegre: SBC/ACN, 2013. p. 188-197.

SILVESTRE, J.C; AGUILERA, V.N. Morte e Luto no Ciberespaço. 2008. Disponível em <http://www.cencib.org/simposioabciber> . Acesso em 20/08/2014.

Projeto de LEI n.º 4.847. Disponível em

[http://www.camara.gov.br/proposicoesWeb/prop\\_mostrarintegra;jsessionid=F813FC880B7DD30A2B70137B8B3C1347.proposicoesWeb2?codteor=1049733&filename=PL+4847/2012](http://www.camara.gov.br/proposicoesWeb/prop_mostrarintegra;jsessionid=F813FC880B7DD30A2B70137B8B3C1347.proposicoesWeb2?codteor=1049733&filename=PL+4847/2012). Acessado em 20/08/2014

Projeto de LEI nº 4099/2012. Disponível em

[http://www.camara.gov.br/proposicoesWeb/prop\\_mostrarintegra?codteor=1004679&filename=PL+4099/2012](http://www.camara.gov.br/proposicoesWeb/prop_mostrarintegra?codteor=1004679&filename=PL+4099/2012). Acessado em 25/09/2014

## **Automação do cálculo de velocidade média e estudo de conceitos de física térmica e mecânica com um barco a vapor**

**Eduardo N. Obara<sup>1</sup>, Devair F. G. Junior<sup>1</sup>, Leonardo P. Souza<sup>1</sup>, Jhony A. Souza<sup>1</sup>, Eliel R. Lima<sup>1</sup>**

<sup>1</sup>IFMT Campus Cáceres. Av. dos Ramires, s/n, Distrito Industrial, 78200-000, Cáceres-MT

{[eduardoobara](mailto:eduardoobara), [leonardodesouza97](mailto:leonardodesouza97), [jhonyconq](mailto:jhonyconq)}@hotmail.com,  
[eliel.regis@cas.ifmt.edu.br](mailto:eliel.regis@cas.ifmt.edu.br)

***Abstract.** This paper describes the work done on thermal physics. This study involves experiments concerning the operation and movement of a steamboat, focusing on the production of an automation system with sensors and software to perform calculations of speed.*

***Resumo.** Este artigo descreve o trabalho realizado sobre física térmica. Este estudo envolve experimentos relativos ao funcionamento e movimento de um barco a vapor, tendo como foco a produção de um sistema de automação com um software e sensores para realizar cálculos da velocidade.*

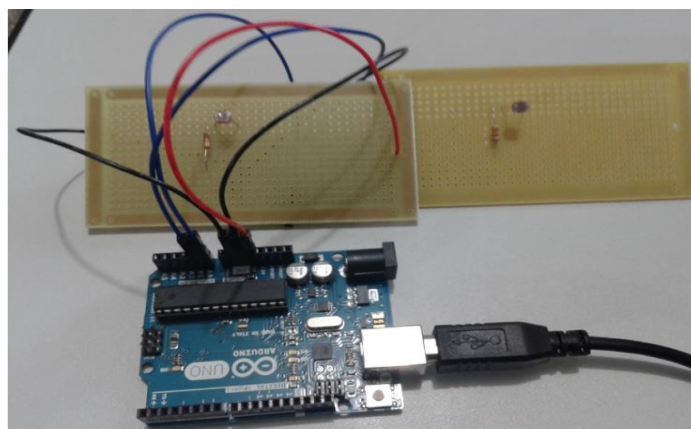
### **Introdução**

O objetivo deste trabalho é construir um sistema automatizado para o estudo de física térmica e mecânica, através da interação entre um software e um sistema de navegação composto por um recipiente com água e um barco a vapor. A comunicação entre o sistema de navegação e o software é realizada por uma placa de automação arduino.

A automação consiste em captar automaticamente os momentos da saída e chegada do barco a vapor, em um percurso no sistema de navegação, transmitir os dados ao computador para, através do software, processar os dados, exibir e armazenar os resultados.

### **Desenvolvimento**

Um software - ou programa de computador - é um conjunto de instruções e comandos que orientam o computador na execução de uma tarefa [Davis, 1991]. O software utilizado neste projeto foi desenvolvido na linguagem de programação Delphi, que é uma linguagem orientada a objeto com uma interface gráfica de desenvolvimento [Sonnino, 2001].



**Figura 1. Sistema com o Arduino e 2 sensores de luminosidade LDR.**

A tecnologia de automação é composta pela placa arduino, que é uma plataforma física de computação, baseada em uma placa microcontroladora de entrada e saída de dados [UFMS, 2012]. Na figura 1 temos o arduino ligado ao computador e conectado a duas placas de circuito, cada uma com um resistor para controlar a energia necessária ao melhor funcionamento do sensor de luminosidade LDR (resistor dependente de luz).

O barco a vapor é uma máquina térmica que transforma a energia térmica em energia mecânica [Hewitt, 2002] e utilizará como meio de navegação um recipiente de mais de dois metros de comprimento, construído especificamente para o experimento.

Em uma lateral do recipiente foram fixados dois sensores de luminosidade LDR que definem o início (ponto A) e o final do percurso (ponto B). Na frente de cada sensor, na outra lateral, foi instalado um laser apontado para o sensor. Os sensores são ligados à placa arduino e enviam sinais para o software: quando o barco passar pelos pontos A (início) e B (final). Ao passar pelo ponto A o cronômetro é iniciado. Ao passar pelo ponto B o cronômetro é finalizado e o software calcula o tempo gasto no percurso entre os pontos A e B e a velocidade média do barco.

A velocidade do barco altera conforme o número de velas e as condições de navegação, sem vento e com vento. É preciso informar a quantidade de velas e se há vento no trajeto. Conforme estabelecida uma relação, é possível ao usuário agrupar os resultados em categorias com a mesma distância de percurso.

## **Resultados e Discussões**

Conforme observado nos testes 1, 2 e 3 (tabela 1), a velocidade média, em um percurso de dois metros, aumenta proporcionalmente ao aumento do número de velas, diminuindo o tempo de deslocamento, o que significa o uso de mais energia para a água em seu interior, maior quantidade de água sai do canudo e exerce uma força maior sobre a água externa, fazendo com que a reação dessa água sobre o barco seja também maior.

No teste 4 (tabela 1), o vento age como uma força contrária a esse movimento, ou seja, como uma força de atrito, diminuindo a velocidade do barco e aumentando o tempo de

percurso.

*Tabela 1: Relação dos testes realizado com o barco a vapor.*

| Nome do Teste | Posição Inicial | Distância Percorrida | Tempo de Deslocamento | Quantidade de Velas | Simulador de Vento está ligado (Sim ou Não) | Velocidade Média |
|---------------|-----------------|----------------------|-----------------------|---------------------|---|------------------|
| Teste 1       | 0m              | 2m                   | 18s                   | 1                   | Não   | 0,11 m/s         |
| Teste 2       | 0m              | 2m                   | 16s                   | 2                   | Não   | 0,12 m/s         |
| Teste 3       | 0m              | 2m                   | 15s                   | 3                   | Não   | 0,13 m/s         |
| Teste 4       | 0m              | 2m                   | 27s                   | 1                   | Sim   | 0,07 m/s         |

As comparações entre os testes são possíveis devido à automação do sistema, que capta o tempo inicial e final de forma automática, tarefa que seria difícil de ser realizada manualmente e com a mesma precisão, podendo ocasionar erros e interferir no resultado.

## Conclusão

O experimento se mostrou eficiente para ajudar no estudo e compreensão dos conceitos de física térmica e mecânica: radiação, condução, convecção, dilatação térmica, vaporização, condensação, velocidade média e forças. A visualização desses conceitos facilita a compreensão dos processos que fazem o barco a vapor se mover.

Com este trabalho foi possível integrar conhecimentos de programação de computador com o estudo de física, produzindo um instrumento que pode tornar mais agradável o estudo de conceitos que, geralmente, são de difícil compreensão para estudantes do ensino médio.

## Referências

- Davis, Willian S. (1991) “Sistemas Operacionais: uma visão sistemática”. Rio de Janeiro: Campus.
- Hewitt, Paul G. (2002) “Física conceitual” 9 ed. Porto Alegre: Bookman.
- Sonnino Bruno. (2001) “Desenvolvendo Aplicações com Delphi 6”. São Paulo: Makron Books.
- UFMS. (2002) Grupo de Robótica. “Introdução ao Arduino”. Disponível em: [http://destacom.ufms.br/mediawiki/images/9/9f/Arduino\\_Destacom.pdf](http://destacom.ufms.br/mediawiki/images/9/9f/Arduino_Destacom.pdf). Acessado em 15/07/2014.

## Automação Residencial: Buscando melhorias na acessibilidade de baixo custo.

Gustavo Pereira da Silva Guedes, Kelven Douglas da Silva, Aline P. Domingos de Souza<sup>1</sup>

<sup>1</sup>UNIRONDON– Centro Universitário Candido Rondon  
Avenida Beira Rio, 3001. Jardim Europa. 78.015-480 – Brazil

gnome\_gustavo@hotmail.com; [kelven.douglas19@gmail.com](mailto:kelven.douglas19@gmail.com);  
[aline.souza@pitagoras.com.br](mailto:aline.souza@pitagoras.com.br)

**Abstract.** *The use of open source hardware and software elements to automation will be shown which is meant to display computing devices such as raspberry-pi, arduino and virtual assistant with artificial intelligence will enable a low cost automation using wireless communication infrastructure for domestic appliances and thus improving accessibility and having safety, comfort and economy.*

**Resumo.** *Será demonstrado a utilização de elementos opensource hardware e software para a domótica cujo tem objetivo de mostrar dispositivos computacionais como raspberry-pi, arduino e assistente virtual com inteligência artificial que possibilitará uma automação de baixo custo usando infraestrutura wireless para a comunicação dos dispositivos domésticos e assim melhorando acessibilidade e tendo segurança, conforto e economia.*

### 1. Introdução

A evolução dos processos automatizados na indústria trouxeram padrões, uniformidade e flexibilidade nos meios de produção. Com o intuito de melhorar a qualidade de vida das pessoas, a automação vem se estendendo do ambiente industrial para o ambiente residencial (FERREIRA, 2010). Mesmo proporcionando muitos benefícios para seus usuários ainda possui seu custo de implantação muito elevado, e um dos motivos desse alto custo é a necessidade de hardwares específicos para o controle residencial. A idealização é propor uma automação de baixo custo tendo acessibilidade, e inteligência artificial, que possa interagir com o usuário por meio de voz, dispositivo móvel ou fixo.

Uma definição mais atual da demótica é a utilização simultânea da eletricidade, da eletrônica e das tecnologias da informação no ambiente residencial, permitindo realizar a sua gestão, local ou remoto e oferecer uma vasta gama de aplicações nas áreas da segurança como conforto, comunicação e gestão de energia (MARIOTONI; ANDRADE,2007).

O papel da demótica é possibilitar o aumento da eficiência e da qualidade de vida do ser humano. Ele está ligado ao uso eficaz da energia e dos demais recursos naturais, onde uma central de controle será conectada aos periféricos de saída e entrada da casa, podendo receber comandos de um computador para realizar as ações

programadas dentro do Micro controlador, isto possibilita um custo muito alto para elaboração.

Visando diminuição dos valores de implantação da automação residencial, surge a possibilidade de utilizar uma categoria de hardware recente no mundo da computação denominada opensource hardware (OSH), ou hardware livre. O objetivo deste trabalho é desenvolver uma residência inteligente totalmente opensource, usando ferramentas de hardware e software livre, permitindo uma adaptação ou modificação em sua plataforma lógica ou física, a ideia é desenvolver uma casa inteligente que aceite todas as necessidades de acessibilidade usando uma assistente virtual tendo algoritmos de inteligência artificial.

Por trás de todo um sistema automatizado existem circuitos lógicos programáveis (CLP) que recebe uma informação, e devolve-a processada consequentemente executando uma rotina. Pensando nisto pode ser utilizado o Arduino ele é um (CLP) opensource de baixo custo com um sistema integrado de desenvolvimento que possibilita a escrita de linhas de comandos na placa (Arduino,2012).

Já o microcomputador Raspberry PI é um computador de placa única desenvolvida pela fundação Raspberry PI. O objetivo principal da fundação é a utilização deste para o uso de estudantes em escolas carentes, porém surgiram n variáveis que podem aproveitar os benefícios oferecidos a um custo inicial de apenas vinte e cinco dólares (\$ 25,00). Possui um processador de 700 Mhz, com uma GPU adequada e uma memória RAM 256 megabytes dando suporte para usar sistema operacional Linux, e com isso diminuindo mais o custo do projeto e totalmente opensource (Rasoverry,2012).

Será utilizado um assistente virtual opensource, dentro do sistema operacional Linux usando o minicomputador Raspberry PI conjugado no Arduino, tendo capacidade de controlar a residência totalmente com todo conforto, economia, segurança e acessibilidade. Assistente virtual é agente de interface representa através de personagem que visam aperfeiçoar o processo de interação entre humanos e computadores. Estes personagens são normalmente associados a algum tipo de mecanismo de inteligência artificial que lhes permitem detectar estímulos externos e responder a estes adequadamente. A inteligência artificial sistematiza e automatiza tarefas intelectuais e, portanto, é potencialmente relevante para qualquer esfera da atividade intelectual humana (RUSSELL; NORVIG, 2004).

O grande problema é a infraestrutura na automação é a ligação de fios, ou seja, a estrutura de cabos, pois é necessários fazer reforma na casa. Usando as ferramentas disponíveis no mercado como o Arduino e Raspberry-pi, não ha necessidade, instalar cabos para fazer troca de informações com aparelhos domésticos, pois tem módulo wireless como exemplo (modulo xbee) que pode fazer as trocas de informações para efetuar a operação desejada sem fio e isso reduz o custo.

A automação com a assistente consegue entrar no conceito de acessibilidade, para o usuário domestico, ela tem a capacidade de interagir com os humanos ao redor da residência, para pode fazer controle na casa dentro da automatização e, outros papeis simples como fazer pesquisa, enviar e-mails permitir conexões de usuários da casa via internet, e muito mais recursos dinâmicos tendo segurança.

### 3. Conclusão

As pessoas com necessidades especiais encontram problemas diários em se tratando de acessibilidade, seja nas ruas bem como na sua própria casa, várias soluções surgem para resolver muitos desses problemas, a maioria dos casos com tecnologia são ineficientes para melhorar a acessibilidade ou tem um custo de implementação muito alto. Foi demonstrado que a aplicação dos elementos open source hardware para domótica apresentou melhores resultados e custo relativamente baixo para a criação de uma residência automatizada para acessibilidade

#### 4. Referências

FERREIRA, Carlos José G. Automação Residencial com Domótica. Disponível em <http://www.artigonal.com/seguranca-da-casa-artigos/automacao-residencial-domotica-369823.html>

MARIOTONI, C. A. e ANDRADE Jr., E. P., Descrição de Sistema de Automação predial Baseados em protocolos PLC Utilizados em Edifícios de Pequeno Porte e Residência, Revista de Automação e Tecnologia de Informação. Volume 1, Numero 1, 2007

Arduino Corp. Arduino Uno 2012. Disponível em:

<http://arduino.cc/en/Main/arduinoBoardUno>

Raspberry Corp. Raspberry PI 2012. Disponível em <http://www.raspberrypi.org>

RUSSEL, Stuart; NORVIG, Peter. Inteligência Artificial. 2 Ed. Rio de Janeiro: Campos, 2004.

## Computação Gráfica na Educação Pública Nacional

Marlon Ribeiro Bordin, Aline Paulino Domingos de Souza <sup>1</sup>

<sup>1</sup>UNIRONDON– Centro Universitário Candido Rondon  
Avenida Beira Rio, 3001. Jardim Europa. 78.015-480 – Brazil

marlow\_bordin@hotmail.com; aline.souza@pitagoras.com.br

**Abstract.** *In this article, it was made an analysis of the Brazilian's public education level, through visiting their places and viewing results of international research. Also, it was conducted an analysis of the government's education software, that will be deployed in public's schools, and starting from these studies we've noted the need of change this bad reality. It was created a software's developing methodology with the purpose to assist the teacher in his ministrations of content, implementing computer graphics, helping to view each module, making the learn more dynamic and fruitful, reaching teaching quality.*

**Resumo.** *Neste artigo, foi efetuado uma análise da realidade educacional nas escolas públicas de ensino fundamental e médio brasileiras através de visitas em seus locais e resultados de pesquisas internacionais. Também realizou análise de um software a ser implantado pelo governo nessas escolas, e, a partir desses estudos, notou-se a necessidade de buscar melhorias na qualidade do ensino nacional. Criou-se uma metodologia para o desenvolvimento de software com finalidade de auxiliar o professor em sua ministração de conteúdo, com a presença da computação gráfica, de modo a ser implantada na visualização de cada módulo, tornando o aprendizado mais dinâmico e proveitoso, gerando um aumento na qualidade de ensino.*

## **1. Introdução**

Na educação pública brasileira, os níveis médio e fundamental apresentam qualidade de ensino insatisfatória, permanecendo abaixo da média, quando comparada com demais países.

A Pearson, empresa de sistemas de aprendizado, publicou em maio de 2014 o segundo relatório avaliativo, no qual o Brasil se encontra na 38ª posição dos 40 países avaliados. O ranking é gerado pelos resultados de testes internacionais (PISA, TIMSS e PIRLS) e pelo índice de alfabetização e aprovação escolar.

O déficit da educação pública brasileira é gerado por diversos fatores, a desvalorização da profissão professor, a ausência de infraestrutura e materiais didáticos, evasão dos alunos, dentre outros.

Este artigo foi desenvolvido objetivando buscar melhorias na qualidade do ensino fundamental e médio públicos do Brasil, demonstrando uma metodologia de desenvolvimento de software em conjunto com a computação gráfica, que poderá ser implantado para prestar auxílio na ministração das aulas dos professores, ampliando o aprendizado e melhorando esta realidade.

## **2. A Proposta de Software Educacional desenvolvida pelo Governo**

Sabemos que existem diversas barreiras na educação nacional, o governo, na tentativa de superá-las, implantará em algumas escolas estaduais de Mato Grosso, um software de apoio à aprendizagem para os alunos, o sistema está sendo verificado pela Secretaria de Educação – SEDUC e será implantado como teste.

Neste software podemos notar a presença de animações em flash, no qual o aluno pode interagir e aprender num ambiente virtual (simulando laboratórios de química, por exemplo), pois esta real infraestrutura é ausente nas escolas.

Esta solução proposta parece viável, entretanto, ao realizar visitas em escolas públicas dos municípios de Cuiabá e Várzea Grande, notei que nos laboratórios de informática utilizados pelos alunos, todas as estações de trabalho utilizam o sistema operacional Linux Educacional 3.0 (Pois três usuários podem utilizar um único microcomputador de modo distinto), onde há incompatibilidade com os sistemas instalados na plataforma Windows (salvo na utilização de emuladores) como é o caso do software do governo. Também a ausência de equipamentos funcionais na maioria das escolas é quase total, o que limitaria a quantidade de alunos que poderão aprender com este sistema.

## **3. Metodologia de Desenvolvimento de Software com Computação gráfica**

Podemos projetar um software que não seja destinado aos alunos individualmente, sem a necessidade de obter vários computadores para seu funcionamento, pois realmente, não existe a quantidade suficiente para os discentes, a infeliz realidade é que a maioria dos equipamentos não funciona. Um sistema no qual não substituirá a importante função do professor, mas sim um software de apoio ao professor na ministração de seus conteúdos.

Seguindo esta metodologia e estes ideais, poderemos desenvolver um programa computacional que será utilizado pelo docente em suas aulas, tornando-as mais dinâmicas e proveitosas, aumentando a absorção de conhecimentos por seus alunos.

Um sistema web pode ser implantado, onde inicialmente, um servidor em

território nacional, hospedará um site/aplicação leve, que permitirá a visualização de animações já processadas de acordo com o conteúdo ministrado. Se o professor estiver ministrando o conteúdo de biologia, por exemplo, artrópodes (seres com patas articuladas), chilopoda (seres com um par de patas por seguimento, por exemplo, a lacraia), através de sub menus a animação gráfica poderá ser encontrada e exibida aos alunos, a presente proposta é demonstrar uma metodologia de desenvolvimento de software juntamente com a computação gráfica, que permitirá a visualização do conteúdo de forma animada e interativa.

Isto enriquecerá a aula ministrada pelo professor, em que, ainda utilizando o exemplo da chilopoda, poderá mostrar aos discentes suas propriedades (alimentação, estrutura interna, etc.) contendo diversos vídeos computacionais para cada uma de suas características, sendo exibidos ao decorrer do conteúdo, permitindo ao aluno não só a ouvi-lo, mas a visualizá-lo e melhor compreendê-lo.

Mas como este sistema poderá ser implantado em escolas com escassez computacional? Os requisitos de infraestrutura tecnológica serão reduzidos apenas a um retroprojetor, um computador, uma caixa de som amplificada e conexão com internet para acessar a aplicação no servidor, tornando o sistema de apoio ao professor capaz de ser implantado mais rapidamente e possivelmente melhorar a qualidade de ensino nas instituições públicas que o adotarão.

#### **4. Resultados**

Ao projetar e demonstrar uma metodologia para o desenvolvimento de software para apoiar a educação pública nacional, problemas no sistema educacional foram descobertos, e tratados de modo prioritário, estudou-se a possibilidade de contorná-los de modo simples e eficaz, construindo e implantando um software que juntamente com a vasta gama da computação gráfica (animações computacionais do conteúdo ministrado pelo docente), tornarão o ensino mais interessante para os alunos, não excluindo o professor deste meio, que será beneficiado por este projeto, pois visa auxiliá-lo em suas aulas.

#### **Referências**

- Joel Edu Sánchez Castro, Fernando Hattori, Felipe Martins dos Santos, Miguel Angel Galarreta Valverde, Instituto de Matemática e Estatística, Universidade de São Paulo. UTILIZAÇÃO DE COMPUTAÇÃO GRÁFICA NA MELHORIA DA EDUCAÇÃO DE CRIANÇAS. Disponível em:  
[HTTP://WWW.IME.USP.BR/~EDU/COMPUGRAFICA/CG\\_FAPESP.PDF](http://www.ime.usp.br/~edu/compugrafica/cg_fapesp.pdf)
- Web Site BBC, “Brasil se distância de média mundial em ranking de educação” (2014), Disponível em:  
[http://www.bbc.co.uk/portuguese/noticias/2014/05/140508\\_brasil\\_educacao\\_ranking\\_dg.shtml](http://www.bbc.co.uk/portuguese/noticias/2014/05/140508_brasil_educacao_ranking_dg.shtml)

# Uma Revisão de Literatura sobre Recursos de Interação Tridimensional para embasar o desenvolvimento de Brinquedo Terapêutico Digital

**Eduardo M. Lemos, Eunice P. dos Santos Nunes, Cristiano Maciel**

Instituto de computação – Universidade Federal de Mato Grosso (UFMT)  
Cuiabá – MT – Brasil

[eunice.ufmt@gmail.com](mailto:eunice.ufmt@gmail.com), [emartinslemos@gmail.com](mailto:emartinslemos@gmail.com), [crismac@gmail.com](mailto:crismac@gmail.com)

***Abstract.** The Therapeutic Toy is an important tool used to simulate hospital procedures. Health professionals generally use physical materials, like dolls or hospital accessories to simulate real situations. As an alternative form of Therapeutic Toy, this paper investigates if Tridimensional Virtual Environment has been used as a Digital Therapeutic Toy to simulate situations experienced by hospitalized children and what interaction resources has been applied.*

***Resumo.** O Brinquedo Terapêutico é uma ferramenta utilizada para simular procedimentos hospitalares. Profissionais da área de saúde em geral, utilizam materiais físicos, como bonecos e acessórios hospitalares para simular situações reais. Como forma alternativa de Brinquedo Terapêutico, este trabalho investiga se o Ambiente Virtual Tridimensional tem sido usado como Brinquedo Terapêutico Digital para simular situações vivenciadas pelas crianças hospitalizadas e quais os recursos de interação vem sendo aplicados.*

## 1. Introdução

O Brinquedo Terapêutico (BT) é uma ferramenta utilizada por profissionais da área de saúde, como forma de simular procedimentos hospitalares. Em geral, o BT é aplicado para explicar à criança sobre o procedimento a que será submetida, buscando prepará-la, ou como forma de descarregar as suas tensões após o processo (FONTES *et al.*, 2010). Esse recurso busca desenvolver a confiança entre a criança hospitalizada e o profissional de saúde, estimula a recuperação e age como exercício físico e psicológico.

De acordo com Silva *et al.* (2010), a hospitalização não é uma tarefa trivial para as crianças, pois carrega consigo não só o seu corpo doente, mas também os seus costumes, sua rotina e sua principal característica: o brincar. No trabalho de Kiche e Almeida (2009), observa-se que há evidências de que a utilização de Brinquedo Terapêutico antes do processo de troca de curativos cirúrgicos reduziu significativamente os índices de dor e tensão muscular das crianças, diminuindo assim a resistência ao procedimento.

A sessão com Brinquedo Terapêutico precisa ser acompanhada por um profissional da área de saúde. Em geral, os profissionais utilizam materiais físicos, como bonecos e acessórios hospitalares para as crianças externarem os seus sentimentos. Porém, além de materiais físicos, há possibilidade de se explorar

Ambientes Virtuais Tridimensionais (AVs 3D), como forma de desenvolver uma versão digital do BT, aqui denominado de Brinquedo Terapêutico Digital (BTD) e que pode também ser aplicado para diferentes propósitos.

O presente trabalho faz parte de um contexto maior que visa a estabelecer um Modelo de Referência para desenvolvimento de Brinquedo Terapêutico Digital. O objetivo principal deste artigo é investigar se AVs 3D têm sido usados para simular situações reais vivenciadas pelas crianças no ambiente hospitalar e quais recursos de interação têm sido aplicados. Para tal, realizou-se uma revisão bibliográfica exploratória sobre o tema, cujos resultados parciais são discutidos a seguir.

## 2. Trabalhos Relacionados

Garzotto e Forfori (2006) desenvolveram um AV colaborativo baseado em hiperhistórias, que busca auxiliar no desenvolvimento psicomotor da criança, uma vez que utiliza de narrativas e histórias cooperativas baseadas em *games* e proporciona a interação com outras crianças. Este tipo de recurso interativo pode auxiliar na assistência à criança hospitalizada, melhorando sua socialização e servindo como uma ferramenta de entretenimento no ambiente hospitalar.

No trabalho de Bickmore *et al.* (2009), os autores desenvolveram um ambiente virtual com “Humanos Virtuais” para representar enfermeiras no ambiente hospitalar, com a finalidade de remover as barreiras de comunicação entre o profissional da saúde e as crianças. Deste modo, as crianças conversam com as enfermeiras virtuais sobre diversas situações hospitalares de maneira descontraída, desenvolvendo a habilidade de comunicação. Além disso, a criança pode interagir com a enfermeira a qualquer momento, reduzindo a sensação de abandono, considerando que em hospitais de baixa renda há carência de profissionais da área de saúde. Os autores realizaram experimentos com crianças hospitalizadas e os resultados evidenciaram que “Humanos Virtuais” podem contribuir para a habilidade de comunicação e auxiliar no tratamento terapêutico.

No estudo apresentado por Akabane *et al.* (2011), os autores desenvolveram o AV 3D colaborativo “*Puchi Planet*” com propósito de oferecer entretenimento com conhecimento. O ambiente simula uma “viagem” ao redor do mundo, utilizando objetos tridimensionais, como globo terrestre, câmera fotográfica, avião e adesivos virtuais. No AV, a criança pode atuar como fotógrafo ou piloto durante a exploração, além de armazenar os dados que considerar relevantes e os locais mais interessantes. Quando a criança permanece longos períodos hospitalizadas, há uma tendência na redução da comunicação e a interação, que muitas vezes, se limita aos membros da família. Além disso, ela deixa de descobrir novos lugares, o que pode desestimular a curiosidade e a aquisição de conhecimento. Dessa forma, o “*Puchi Planet*” busca estimular a interação da criança com outras pessoas, por meio de um ambiente colaborativo, além de favorecer a exploração de lugares desconhecidos.

Um AV 3D baseado em *games* é apresentado por González, Collazos e González (2012), que envolve fatores como diversão, competição e diferentes tipos de comportamentos, com o propósito de melhorar a motivação, a comunicação e o trabalho em equipe das crianças hospitalizadas. Devido à doença e ao seu tratamento, as crianças estão suscetíveis a desenvolver *stress* e ansiedade, e a falta de interação com familiares e amigos pode diminuir a eficiência do tratamento médico, causar traumas psicológicos e diminuir a motivação para aprendizagem. Os autores analisaram os fatores humanos

referentes ao AV e concluíram que os resultados foram satisfatórios, reduziu o isolamento das crianças no ambiente hospitalar e aumentou a motivação para aprender. Deste modo, AVs 3D baseado em *games* podem colaborar na redução do nível de *stress*.

### 3. Conclusões

Considerando a revisão apresentada, observa-se na literatura a concepção de AVs 3D aplicados na assistência à criança hospitalizada. Os principais recursos de interação identificados estão baseados em *games*, cenários colaborativos e “Humanos Virtuais”. Observa-se também que as iniciativas são aplicadas em situações específicas, que buscam explorar a comunicação, a socialização, o entretenimento, a exploração de lugares desconhecidos ou como forma de estimular a aprendizagem.

No entanto, com base nesta consulta inicial, não foram encontradas iniciativas que, de fato, simulem situações reais vivenciadas pelas crianças no ambiente hospitalar, como forma de prepará-las antes do procedimento médico ou com propósito de descarregar a tensão após o processo. Contudo, os recursos de interação identificados podem ser um norte para a continuidade desta pesquisa, cujo próximo passo é realizar uma Revisão Sistemática (RS) na literatura especializada, buscando investigar pesquisas consolidadas que aplicam Brinquedo Terapêutico simulando cenários clínicos da vida real e pesquisas que concebem ou aplicam Brinquedo Terapêutico Digital.

### Referências

- Akabane, S., Leu, J., Iwadate, H., Inakage, M., Nakayama, S., Chang, E., Furukawa, S. (2011). "Puchi Planet : A Tangible Interface Design for Hospitalized Children", In: CHI'11 Extended abstracts on Human Factors in Computing Systems, New York, USA., p. 1345-1350.
- Bickmore, T. W., Pfeifer, L. M., & Jack, B. W. (2009). "Taking the Time to Care : Empowering Low Health Literacy Hospital Patients with Virtual Nurse Agents", In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, New York, USA., p. 1265-1274.
- Fontes, C., Mondini, C., Moraes, M., Bacheга, M., Maximino, N. (2010). "Utilização do Brinquedo Terapêutico na Assistência à Criança Hospitalizada", Rev. bras. educ. espec. v. 16, n. 1, p. 95-106.
- Garzotto, F., & Forfori, M. (2006). "Hyperstories and Social Interaction in 2D and 3D Edutainment Spaces for Children", In: Proceedings of the Seventeenth Conference on Hypertext and Hypermedia, New York, USA., p. 57-68.
- González, C. S., Collazos, C., & González, J. L. (2012). "The importance of human factors to enhance the user experience in videogames", In: Computers in Education (SIIE), Andorra La Vella, Andorra., p. 1-4.
- Kiche, M., Almeida, F. (2009). "Brinquedo terapêutico : estratégia de alívio da dor e tensão durante o curativo cirúrgico em crianças", Acta Paul. Enferm. v. 22, n. 2, p. 125-130.
- Silva, S., Jesus, I., Santos, R. (2010) “Humanização em Pediatria: O Brinquedo Terapêutico Como recurso na Assistência de Enfermagem à Criança Hospitalizada”. Revista Brasileira de Medicina, LILACS, 2010, v. 46, n<sup>o</sup>. 3, p. 101- 104.

# Avaliação de métodos de combinação de classificadores em um sistema de auxílio ao diagnóstico de pneumonia infantil

Mathias Fassini Mantelli<sup>1</sup>, Rafael Teixeira Sousa<sup>1</sup>

<sup>1</sup>Instituto de Ciências Exatas e da Terra – Universidade Federal de Mato Grosso (UFMT)  
Barra do Garças – MT – Brazil

mathiasfassini@gmail.com, rafaelts777@gmail.br

**Abstract.** *In this paper the results of the combination methods Bagging and Boosting which combine the classifiers K-Nearest Neighbors (KNN) and Naive Bayes (NB) with the goal of improving the diagnosis of childhood pneumonia by analysis of chest radiographs of children are analyzed.*

**Resumo.** *Neste trabalho serão analisados os resultados dos métodos de combinação Bagging e Boosting, que combinam os classificadores K-Nearest Neighbors (KNN) e Naive Bayes (NB) com o objetivo de melhorar o diagnóstico da pneumonia infantil por meio da análise de radiografias torácicas de crianças.*

## 1. Introdução

Pneumonia infantil é uma doença epidêmica caracterizada por infecção respiratória aguda baixa, geralmente causada por vírus, bactérias ou fungos. De acordo com a [World Health Organization 2013], a pneumonia lidera as causas de mortes infantis em todo o mundo, matando mais de 1,1 milhões de crianças com até cinco anos de idade em todos os anos. Este número é maior que a mortalidade de várias outras doenças juntas, como a AIDS, malária e tuberculose.

Segundo a [World Health Organization 2001], última publicação da WHO a respeito do diagnóstico de pneumonia infantil, a radiografia torácica é o melhor método e também o mais aceito para diagnosticar a pneumonia. Porém, erros são comuns na interpretação das mesmas, devido às variações do inter-observador [Young e Marrie 1994]. Esses erros tornam-se uma motivação para o uso de computadores para melhorar a velocidade e a acurácia do processo de diagnóstico.

Neste trabalho utilizaremos os recursos e conjuntos de dados utilizados em estudos anteriores [Macedo e Oliveira 2012], [Sousa et al. 2013], que resultou em um programa completo para o diagnóstico de pneumonia. Este programa servirá como uma segunda opinião sobre o diagnóstico [Doi et al. 1999].

Este trabalho tem como objetivo uma análise do desempenho de dois métodos da combinação de classificadores, para melhorar a acurácia e descobrir o melhor método de combinação para o diagnóstico de pneumonia infantil.

### A. Métodos de Combinação selecionados

Neste trabalho nós aplicamos dois métodos de combinação e dois classificadores: como métodos, citamos o *Bagging* e *Boosting*; como classificadores, citamos o *K-Nearest*

*Neighbors (KNN)* e *Naive Bayes (NB)*. Os classificadores foram selecionados segundo estudos anteriores que mostraram seu desempenho superior no problema em questão [Sousa et al. 2013].

*Bagging* utiliza o método de votação simples e robusto. Utilizando múltiplas versões de um conjunto de treinamento, cada versão é criada selecionando aleatoriamente  $n' < n$  amostras do conjunto de treinamento  $D$ , com reposição. Essas versões são utilizadas para treinar diferentes classificadores e a decisão final é baseada no voto de cada um desses classificadores [Carlos et al. 2005].

O procedimento *Boosting* consiste em aplicar, sequencialmente, uma regra de classificação qualquer (chamada de classificador base) a versões iterativamente ponderadas da amostra de treinamento. Em cada iteração o algoritmo atribui pesos maiores às observações incorretamente classificadas na iteração anterior. A regra final de classificação é obtida através de uma combinação linear dos classificadores construídos em cada iteração, produzindo um comitê de classificação [Marcelo 2007].

## 2. Metodologia

O conjunto de imagens usadas no programa contém 156 unidades, em escala de cinza de 8 bits, obtidas com uma câmera digital, que capturou radiografias torácicas em uma resolução de 1024 x 768 pixels. Destas imagens, 78 se mostraram positivas para pneumonia infantil e 78 não. Estas imagens foram analisadas por dois radiologistas treinados de acordo com os parâmetros da *WHO* [Cherian et al. 2005], que produziram o diagnóstico real para testar os classificadores.

Foram utilizados 14 descritores baseados em textura, sendo extraídos em nove subespaços obtidos com a *Wavelet de Haar*, após a extração foi feita a redução da dimensionalidade com o KPCA [Sousa et al. 2013].

Nos testes com o método *Bagging*, treinamos 10 classificadores, selecionando 70% do conjunto de teste para cada classificador. O método de *Boosting* foi aplicado com o algoritmo *AdaBoost*, também com 10 classificadores.

Ambos os métodos de combinação foram executados com os dois classificadores selecionados (*KNN* e *NB*) com 10 execuções usando validação cruzada com 10 *folds*.

## 3. Resultados

Os resultados dos testes estão na Tabela 1, com os dados dos classificadores executados de forma simples (sem combinação) e com os métodos de combinações, avaliados segundo a taxa de acerto.

**Tabela 1. Resultados dos classificadores**

| Método   | Classificador |       |
|----------|---------------|-------|
|          | NB            | KNN   |
| Simple   | 95%           | 90%   |
| Bagging  | 96,8%         | 88,5% |
| Boosting | 97,4%         | 87%   |

Os resultados obtidos mostram uma leve melhora no classificador *NB* e uma piora no *KNN*. O *KNN* teve uma piora quando aplicado aos métodos *Bagging* e *Boosting* em relação ao resultado obtido no classificador individual.

#### 4. Conclusão

O método *Boosting* se mostrou eficiente aplicado ao classificador *NB* e ineficiente ao classificador *KNN*. Quando este método foi aplicado ao *NB*, houve uma melhora em relação a *NB* individual (2,4%) e em relação ao método *KNN* (0,6%).

A respeito do método *Bagging*, todos os resultados foram inferiores ao resultado obtido pelos classificadores individuais. Portanto, conclui-se que o melhor nível de acerto foi obtido através da combinação do método *Boosting* com o classificador *NB*.

Acredita-se que o avanço do classificador *NB* com o método *Boosting* se deve ao tratamento dos erros durante as iterações, pois alguns estudos mostram que este fator pode levar o classificador a desconsiderar atributos irrelevantes [Elkan 1997].

Futuramente devemos testar o uso de outros algoritmos de combinação para melhorar os resultados.

#### References

- Carlos, N. S. Jr., Celso, A. A. K. e Alessandro, L. K. (2005) “Classificação Automática de Gêneros Musicais Utilizando Métodos de ‘*Bagging*’ e ‘*Boosting*’ ”.
- Cherian, T., Mulholland, E. K., Carlin, J. B., Ostensen, H., Amin, R., Campo, M. D., ... e Steinhoff, M. C. (2005) “Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies.” *Bulletin of the World Health Organization*, 83(5), p. 353-359.
- Doi, K., MacMahon, H., Katsuragawa, S., Nishikawa, R. M. e Jiang Y. (1999) “Computer-aided diagnosis in radiology: potential and pitfalls.”, no *European Journal of Radiology*, vol. 31, no. 2, p. 97 – 109.
- Elkan, C. (1997) “Boosting and Naïve Bayesian Learning”.
- Macedo, S. O. e Oliveira, L. L. G. (2012) “Desenvolvimento de um sistema de auxílio ao diagnóstico de pneumonia na infância utilizando visão computacional”, no *Workshop de Visão Computacional*.
- Marcelo, R. P. F. (2007) “Análise Discriminante Clássica e de Núcleo: avaliações e algumas contribuições relativas aos métodos boosting e bootstrap”, p. 64.
- Sousa, R. T., Marques, O., Soares, F. A. A. M. N., Sene, I. I. G., Oliveira, L. L. G., e Spoto, E. S. (2013) “Comparative performance analysis of machine learning classifiers in detection of childhood pneumonia using chest radiographs”, no *ICCS*, p. 2579–2582.
- World Health Organization (2013) “Pneumonia, ficha n° 331”.
- World Health Organization: Department of Vaccines and Biologicals (2001) “Standardization of interpretation of chest radiographs of the diagnosis of pneumonia in children”.
- Young, M. e Marrie, T. J. (1994) “Interobserver variability in the interpretation of chest roentgenograms of patients with possible pneumonia”, No *Arch Intern Med*, vol. 154, p. 2729–2732.

## E-con: uma proposta para modernização do Procon - MT

**Ketelem Lemos de Campos, Bruno Honorato de Oliveira, André Luiz Miranda,  
Cleder Almeida, Cristiano Maciel**

<sup>1</sup>Instituto de Computação, Universidade Federal de Mato Grosso (UFMT)  
Av. Fernando Correa da Costa, 2367, Boa Esperança - Cuiabá - MT – Brasil

***Abstract.** E-con offers consumers conduct their claims for injuries to the consumer law and check them online. In addition to the query of the name for official Credit Protection, and report the situation. This article demonstrates a proposal for implementation of the E-con to PROCON of the Mato Grosso. This proposal is not restricted to simple automation of processes and provision of public services, but rather on the transformation of the way the Government, using new technologies, reaches the objectives for the fulfilment of the role of the state.*

***Resumo.** O E-con propõe ao consumidor realizar suas reclamações por lesões ao direito de consumidor e consultá-las online. Além de fazer a consulta do nome junto aos Órgãos de Proteção ao Crédito, e informar a situação. Este artigo demonstra uma proposta de implementação do E-con para o PROCON de Mato Grosso. Esta proposta não se restringe a simples automação dos processos e disponibilização de serviços públicos, mas sim na transformação da maneira com que o governo, utilizando-se das novas tecnologias, atinge os objetivos para o cumprimento do papel do estado.*

### 1. Introdução

Atualmente no Brasil, o órgão do SINDEC<sup>1</sup> que rege todos os processos dos PROCON's estaduais e, nos estados, poucas são as cidades que possuem seu ponto de atendimento. Em Mato Grosso, por exemplo, são raras as cidades que podem atender o consumidor diretamente. Segundo o site do Procon-MT, em 2013, foram realizados 16.690 registros de reclamações. E por meio da observação e conversa informal com alguns funcionários, para cada atendimento a ser realizado leva-se de 10 a 30 minutos de espera para ser atendido e 40 a 60 minutos no atendimento, totalizando de 50 a 90 minutos, sem contar, é claro, o tempo gasto para chegar ao Órgão. Diante disso, surge a ideia de implementação desta proposta. A ferramenta proposta é um meio de suprir essa necessidade e poderá ser entendida como uma forma de modernização do estado para uso das novas tecnologias para a prestação de serviços públicos, transformando o modo com que o governo interage com os cidadãos.

---

<sup>1</sup> <http://portal.mj.gov.br/sindec>

## 2. Metodologia

O E-con é uma proposta em pesquisa que utiliza métodos e técnicas da Engenharia de Software. Em sua construção está sendo utilizado o Modelo Cascata, pois segundo Sommerville (2011) é o mais usado em projetos de engenharia que há necessidade da validação de todas as suas etapas, tornando este modelo o mais viável, uma vez que o sistema terá que receber uma grande quantidade de dados e deve atender em sua totalidade as necessidades dos operadores, possibilitando a interface e a comunicação com sistemas já existentes. Após a verificação da quantidade de fluxos relacionados as reclamações cadastradas anualmente, disponibilizada do site do Procon-MT, percebeu-se que a solução precisa ser bem discutida e analisada antes de ser implementada, o que é uma características desse modelo.

Para a fase de análise foram elencadas técnicas de elicitação como entrevistas e questionários, uma vez que foi necessário elucidar as principais dificuldades e reclamações dos consumidores lesionados do PROCON, e também facilitar o atendimento com relação aos funcionários em solucionar as dificuldades e encontrar as facilidades no atendimento dos consumidores lesionados. Para detalhamento dos requisitos foram desenvolvidos protótipos em alta e baixa fidelidade. Para implementação do aplicativo web E-con sugere-se as linguagens PHP e HTML.

## 3. Ferramentas Encontradas

Na pesquisa foi apontado o sistema de reclamações online nos sites do PROCON dos Estados do Rio de Janeiro<sup>2</sup>, Minas Gerais<sup>3</sup>, Pernambuco<sup>4</sup> e São Paulo. Recentemente o PROCON do Rio de Janeiro apresentou um aplicativo móvel para esta função, todavia os recursos apresentados são similares ao do site deste Estado. O Estado de São Paulo também possui um atendimento eletrônico, porém, os cadastros devem ser realizados por meio de cartas que são baixadas no site da instituição. As cartas devem ser preenchidas e enviadas com os documentos necessários conforme o órgão determina. Hoje, o Estado de Mato Grosso utiliza esse mesmo mecanismo de cartas, porém, esse sistema contribui com o desperdício de tempo no cadastro da reclamação.

## 3. Proposta de Implementação

Neste item pretende-se discutir como a solução tem sido pensada e especificada. O E-con terá uma página inicial com as informações básicas para o consumidor orientar-se, a respeito do que quer fazer. Nela irá conter uma parte de notícias com atualizações diárias de informações de leis que foram alteradas ou mudanças no código de defesa, além de informações pertinentes ao consumidor. Na lateral esquerda haverá um espaço para downloads de arquivos que irão instruir o consumidor a realizar a sua reclamação, assim como, um chat para tirar dúvidas com um atendente online. No E-con também estará disponível um link para acesso à página de reclamações. Para realizar a reclamação cria-se um usuário por meio do login via CPF deste consumidor e senha. Feito o login estará disponível o espaço para realizar a reclamação, consultar os

---

<sup>2</sup> <http://www.procon.rj.gov.br/index.php>

<sup>3</sup> <https://aplicacao.mpmg.mp.br/ouvidoria/cidadao/aceso.do?idOuvidoria=14>

<sup>4</sup> <http://www.procon.pe.gov.br>

processos já cadastrados e realizar a verificação do seu nome junto aos Órgãos de Proteção ao Crédito. Havendo mudanças no status da reclamação, será enviado um alerta para o consumidor (e-mail ou sms). Na página inicial haverá um campo para pesquisa das empresas que receberam reclamações, apresentando quantas destas reclamações foram solucionadas e quantas ainda aguardam uma solução.

O E-con se difere de outros sistemas por apresentar não só um espaço para cadastramento de informações pessoais e acompanhamento mas por possibilitar uma comunicação mais direta entre o atendente e o consumidor a partir do momento em que são emitidos alertas, sendo e-mails ou SMS para o consumidor poder acompanhar a sua reclamação; um chat para falar diretamente com o atendente, e um espaço com tiradúvidas que possibilitará a um consumidor inexperiente receber auxílio na hora de cadastrar a sua reclamação. Logo, o E-con demonstra uma proposta feita para auxiliar ao consumidor realizar a sua reclamação com eficiência e rapidez em seu atendimento, sem precisar ocupar um grande período de tempo e esforço, e a reclamação poderá ser feita de qualquer lugar ou dispositivo que contenha acesso a internet. Protótipos das interfaces são apresentados a seguir.



Figura1: Protótipo da Página Inicial do E-con

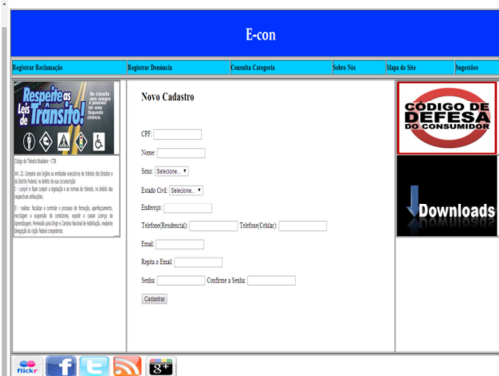


Figura2: Protótipo de cadastro do E-con

#### 4. Conclusão

O E-con trará a facilidade do acesso à informação, a conscientização e a busca pelos direitos do consumidor, possibilitando uma maior interação desses com o órgão regulador. Para que o E-con se torne realidade será necessário o acesso aos dados do PROCON - MT, acesso ao banco de dados dos Órgãos de Proteção ao Crédito, assim como ferramentas de modelagem e construção do código. Todavia, para não termos problemas no acesso indevido a informações e possíveis vazamentos de dados, causando problemas à privacidade do consumidor, são necessárias medidas de segurança, como a criptografia de dados.

#### Referências

SOMMERVILLE, I. Engenharia de Software: 9ª. Edição. São Paulo: Person, 2011.

ALEXANDRIDIS, G. Cadastro de Reclamação Fundamentadas contra Fornecedores - São Paulo : 2014. Disponível em:

<http://georgiosalexandridis.Jusbrasil.com.br/artigos/128429354/cadastro-de-reclamacoes-fundamentadas-contra-fornecedores?ref=home> Acesso em 3 de julho de 2014

Procon MT . Disponível em: [www.procon.mt.gov.br/](http://www.procon.mt.gov.br/) Acesso em 3 de julho de 2014

# Automação Residencial por comando de voz em Delphi 7

**Thassiany B. Pires, Vinicius B. Silva, Fabiano G. Rocha**

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso Campus Cáceres  
Av. dos Ramires, s/n – 78.200-000 – Distrito Industrial – Cáceres – MT – Brasil  
{thassianybpirez, end.viniciussilva, eng.fabianorochoa}@gmail.com

**Abstract.** *This article describes the activities developed in creating a home automation software by voice command. The software in conjunction with a controller card allows you to manage electronic equipment from a residence, enabling to set the date and time the equipment is turned on and off.*

**Resumo.** *Este artigo descreve as atividades desenvolvidas na criação de um software de automação residencial por comando de voz. O software em conjunto com uma placa controladora permite administrar equipamentos eletroeletrônicos de uma residência, permitindo programar a data e hora em que o equipamento será ligado e desligado.*

## Introdução

Com as inovações no mundo tecnológico surgiu a necessidade de automatizar equipamentos eletroeletrônicos e, nesse contexto, também surge a necessidade de desenvolver software para administrar esses equipamentos.

Um problema frequente no dia a dia das pessoas é quando à necessidade de administrar (ligar/desligar) equipamentos eletrônicos, em dias e horários em que o usuário não esteja em casa, como por exemplo, ligar lâmpadas no período noturno, acionar um motor para irrigação de jardim, entre outros. Nessa linha de pensamento, pode-se relacionar a situação problema em que a pessoa com deficiência motora possui dificuldade ao acionar equipamentos eletrônicos.

O software proposto em conjunto com a placa controladora vem a sanar tais situações problema, pois permite o acionamento dos equipamentos a distância ou de maneira presencial aceitando comandos por voz e, conseqüentemente, melhorando a acessibilidade do usuário e sua experiência com o uso do software. Este software foi desenvolvido na linguagem Delphi para controle e automação de sistemas e, por meio de uma interface simples e de fácil manuseio, permite ao usuário programar o acionamento de até oito dispositivos eletrônicos.

## Abordagem teórica

O software está em desenvolvimento utilizando a linguagem de programação Delphi 7. O Delphi é um ambiente e uma linguagem de programação usado para o desenvolvimento de aplicações diversas, compatível com o sistema operacional Windows, é uma linguagem fortemente estruturada que permite não só haver menos erros, mas também que os que há sejam mais fáceis de encontrar e resolver.

Um dos principais recursos do Delphi é o acesso ao banco de dados, que armazena as informações salvas pelo software e pode ser acessado através de diversas tecnologias disponíveis tais como: *BDE (Borland Database Engine)*, *dbExpress (Database Express)* e entre outras. Por meio da interface do software, que foi projetada com os recursos

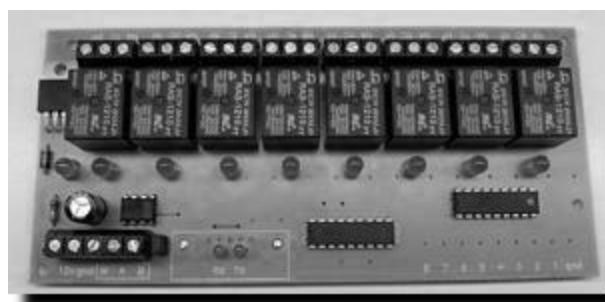
disponíveis no Delphi e o fácil acesso as tabelas do banco de dados, o usuário pode programar o acionamento dos dispositivos.

Este comando é então convertido em um sinal elétrico que é enviado para a placa controladora que determina o controle automático do dispositivo a ela conectada. Na Figura 1 pode-se observar uma das telas do software desenvolvido em que o usuário pode efetuar o gerenciamento dos equipamentos eletroeletrônicos definindo previamente as datas e os horários de acionamento.

| Porta   | Dispositivo       | Hora/Ligar | Data/Ligar | Hora/Desligar | Data/Desligar |
|---------|-------------------|------------|------------|---------------|---------------|
| Porta 1 | Lâmpadas externas | 16:30:00   | 25/10/2014 | 16:31:00      | 25/10/2013    |
| Porta 2 | Lâmpadas externas | 01:37:00   | 25/10/2013 | 01:37:00      | 25/10/2013    |
| Porta 3 | Lâmpadas externas | 01:37:00   | 25/10/2013 | 01:37:00      | 25/10/2013    |
| Porta 4 | Lâmpadas externas | 16:12:00   | 25/10/2013 | 16:13:00      | 25/10/2013    |
| Porta 5 | MICROONDAS        | 01:37:00   | 25/10/2013 | 01:37:00      | 25/10/2013    |
| Porta 6 |                   | 01:37:00   | 25/10/2013 | 01:37:00      | 25/10/2013    |
| Porta 7 | ventilador        | 16:12:00   | 25/10/2013 | 16:13:00      | 25/10/2013    |
| Porta 8 |                   | 02:47:00   | 25/10/2013 | 02:48:00      | 00:12:15      |

**Figura 1: Interface do sistema diário. Fonte: Arquivo pessoal.**

A placa controladora fabricada pela I/O Robotics é conectada à rede elétrica e aos aparelhos eletrônicos, como eletrodomésticos ou lâmpadas. Cada equipamento ligado à placa é conectado a um dos oito relês do circuito da placa. Pode-se observar que na tela ilustrada na Figura 1, o campo *Porta* corresponde ao acionamento de um determinado relê, como por exemplo: a Porta 7 refere-se ao equipamento eletrônico “Ventilador” que está conectado ao relê 7 da placa controladora ilustrada na Figura 2.



**Figura 2: Placa controladora com oito portas. Fone: I/O Robotics.**

No projeto, cada porta é identificada pelo computador através de um conjunto de números. As oito portas disponíveis são totalmente independentes, ou seja, aceitam diferentes comandos simultaneamente. O sistema de automação residencial pode ser controlado por comandos obtidos via teclado, mouse e por comando de voz. Os comandos de voz utilizados no sistema serão pré-definidos e podem ser consultados no menu “Ajuda” para facilitar o manuseio do programa pelo usuário.

## Desenvolvimento

As telas do software que foram desenvolvidas são de fácil uso e acessibilidade. A tela inicial possui o formulário de *login*, em que o usuário se identifica para usar o programa e,

uma vez autenticado, o usuário pode acessar os outros formulários por meio de um menu localizado na parte superior do formulário principal. O software em desenvolvimento utiliza como base o Delphi e seus recursos visuais, como os componentes *LookupComboBox* presente nas telas “Diário” e “Antirroubo”.

O comando de voz viabiliza o uso do software por pessoas com necessidades especiais de locomoção, melhorando a experiência do usuário na gerência do uso dos dispositivos eletrônicos. Para o desenvolvimento do sistema de reconhecimento de voz foi utilizado o Microsoft Speech API. Trata-se de um recurso presente no Windows que disponibiliza as funções de reconhecimento e síntese de voz que estão em implementação no sistema por meio dos componentes que implementam as rotinas que permitem o acesso as funções de reconhecimento de voz.

A comunicação entre a placa controladora e o software é feita utilizando uma biblioteca de código aberto chamada *Comport*, que fornece acesso a recursos básicos de comunicação via porta serial, entre eles leitura e escrita de dados.

O sistema composto pelo computador, o software e a placa controladora, permite o controle de até oito equipamentos, inclusive de voltagens diferentes, de 0 a 220 Volts, ligando e desligando em horários predefinidos, conforme a necessidade do usuário.

## Resultados

Obtivemos sucesso nos testes realizados com a interação do software e a placa controladora, todas as portas responderam corretamente. Os testes com o comando de voz geraram alguns erros que estão sendo corrigidos, somente alguns comandos para controlar as opções de abrir e fechar as telas foram cadastrados como por exemplo: “Sistemas”, e “Abrir Diário”. Notamos que o reconhecimento de voz não funciona de perfeitamente em locais com barulhos, pois interfere no reconhecimento do comando para executar a função dita.

## Conclusão

O protótipo do software grava e acessa os dados definidos pelo usuário, permite controlar equipamentos com voltagens diferentes, em horários e datas pré-definidas ou aleatoriamente. Todas as portas estão funcionando nos horários e datas pré-definidos pelo usuário, os comandos de voz que já foram cadastrados também funcionam corretamente, independente do timbre de voz da pessoa.

Com as pesquisas realizadas foi possível contextualizar as diferentes etapas do processo de desenvolvimento de software, agregar conhecimento sobre automação residencial, bem como entender o funcionamento da placa de automação utilizada e o componente que possibilita o comando de voz.

## Agradecimento

Os autores agradecem a PROPES/IFMT (Edital Propes/IFMT 021/2014) pelo apoio.

## Referências

- DATE, C. (2004) "**Introdução a Sistemas de Banco de Dados**", 8 ed, Rio de Janeiro.
- CANTU, Marco (2002) "**Dominando o Delphi 6 A Bíblia**", São Paulo, Makron Book.
- Alphatron. (2013) "**Automação Residencial**", <http://www.automacaoresidencial.com.br>, Junho.
- I/O Robotics. (2013) "**Placa de 8 relês Multipoint**", [http://www.irobotics.com/placa\\_de\\_8\\_reles\\_multipoint.html](http://www.irobotics.com/placa_de_8_reles_multipoint.html), Junho.

## Motores de Passo: estudo e aplicabilidade em uma maquete interativa controlada em tempo real

Marcos Danillo S. Lau, João Pedro A. S. Pereira, Leandro J. Ferrarezzi,  
Ricardo S. Oliveira, Fabiano G. Rocha

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso Campus Cáceres  
Av. dos Ramires, s/n – 78.200-000 – Distrito Industrial – Cáceres – MT – Brasil

{marcos.danillo1807, leandrojlf, ricardooliveira9819}@gmail.com,  
joaopedrol118@outlook.com, fabiano.rocha@cas.ifmt.edu.br

**Abstract.** *With the interaction between software and hardware, it arises the possibility of controlling electronic devices, such as the pass engines. In this composition a computer software accomplishes the control and automation of pass engines having its applicability in a model controlled in real time. The pass engine is connected to the computer via parallel slot and the software makes electric signals determining the action to be executed by the pass engines in the model.*

**Resumo.** *Com a interação entre software e hardware surge a possibilidade de controlar dispositivos eletrônicos, tais como os motores de passo. Neste trabalho um software de computador efetua o controle e a automatização de motores de passo tendo sua aplicabilidade em uma maquete controlada em tempo real. O motor de passo é conectado ao computador via porta paralela e o software emite sinais elétricos determinando a ação a ser executada pelos motores de passo na maquete.*

### Introdução

A primeira impressão é a que fica e na busca pelo sucesso do empreendimento ou até mesmo de uma simples apresentação de um projeto, se destaca quem é capaz de atrair os clientes e potenciais investidores. Nesse sentido, temos que levar em consideração a interação do cliente com o produto, gerando motivação e atenção. Neste trabalho buscamos fazer a comunicação entre hardware e software com o objetivo de controlar motores de passo via porta paralela do computador. Essa interação entre software e hardware torna possível a análise do comportamento e aplicabilidade dos motores de passo no controle e automatização de uma maquete em tempo real.

Tendo por base projetos já desenvolvidos, buscou-se um software que torne possível controlar motores de passo via porta paralela permitindo o seu uso na automatização de uma maquete. Com o software de computador e os motores de passo idealizamos a maquete, de maneira que seja possível aproximar o usuário de uma experiência diferente das outras já existentes, tal como observar a movimentação de elementos de uma maquete podendo ser composta por: pontes elevadas, movimentação de carros, abertura de portões, guindastes, elevadores, semáforos e até mesmo exercer tarefas simples que fazemos em nosso cotidiano.

Esta mesma linha de pensamento pode ser em outros contextos de informática, tais como na construção de robôs, controle de câmeras ou quaisquer outros dispositivos que possam ser controlados através de motores de passo via software de computador. Esta linha de projeto envolve diversos conhecimentos, tais como: algoritmos, linguagem de programação, eletrônica e uma boa dose de criatividade na construção da maquete.

### Abordagem teórica

Dentre os conteúdos teóricos necessários para o projeto, é preciso compreender o funcionamento do motor de passo, a comunicação via porta paralela, bem como a manipulação adequada de bits e operadores binários AND, OR e de deslocamento.

Os motores de passos são dispositivos mecânicos eletromagnéticos que por podem ser controlados por hardware específico ou através de software. Na Figura 1 e na figura 2 pode-se observar o motor de passo que será utilizado na geração da maquete automatizada. Por meio do software de computador o usuário determina qual a ação que o motor deve executar, qual o número de rotações e em qual direção ele deve ser acionado. A comunicação entre o computador e o motor de passo é realizada por meio da emissão de sinais elétricos pela porta paralela do PC.



Figura 1: Vista do motor passo e circuito controlador. Figura 2: Vista das bobinas do motor.

Toda a lógica de controle dos motores de passo é de responsabilidade do software podendo controlar cada motor de passo independentemente, de forma manual ou automática.

### Desenvolvimento do projeto

O software utilizado está escrito em linguagem C++ e compilado no Borland C++ Builder versão 6.0. Com este software foi possível esquematizar as funções que os motores de passo terão de executar, permitindo a interação do usuário com o software por meio de sua interface gráfica e caixas de mensagem. Dentre as funcionalidades que o software oferece, é possível ligar e desligar, e controlar o giro de motores (direita/esquerda). Na Figura 3 pode-se observar uma das telas do programa utilizado para o controle dos motores de passo.

Com o software funcionando em conjunto com os motores de passo, o próximo passo do projeto consiste na montagem da maquete. De forma simples, a maquete será montada utilizando folha de isopor para a base e as miniaturas da casa serão feitas de papéis especiais para modelagem de objetos.

Nos circuitos serão utilizados fios de cobre para conectar cada motor de passo aos CI (Circuito Integrado) que darão possibilidade para o controle de até oito motores de passo, tendo todos os motores conectados e alimentados via porta paralela. A

maquete será criada com a adição de vários motores de passo em posições estratégicas, com o objetivo de mostrar o máximo de movimento, esses movimentos serão controlados por um computador através de um circuito elétrico.

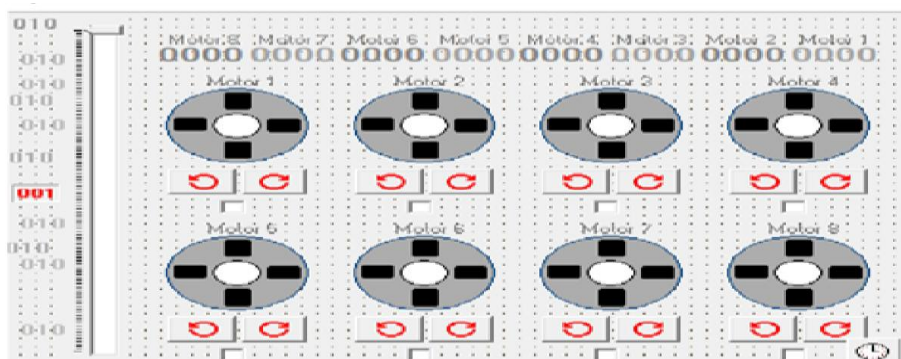


Figura 3: Exemplo de tela do programa.

## Trabalhos Futuros

Tendo em vista os bons resultados já observados com a interação entre o hardware (motores de passo) e o software, e tendo por objetivo proporcionar ao usuário uma melhor experiência com a maquete automatizada, em etapas futuras do projeto buscar-se-á efetuar o controle dos dispositivos eletrônicos via placas Arduino.

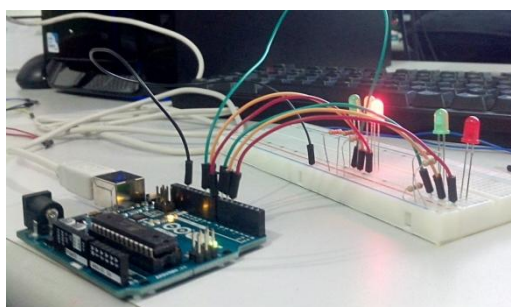


Figura 4: Circuito de um semáforo controlado pela placa Arduino.

Este upgrade no projeto vem a possibilitar uma maior gama de controle e automação da maquete, pois será possível criar de uma maneira mais simples a interface entre um motor de passo e o Arduino por meio da utilização do chip ULN2003A. O ambiente de programação Arduino possui a biblioteca *Stepper* que disponibiliza algumas funções para controlar o motor de passo. Há diversas outras possibilidades de automação com a integração das placas Arduino, tais como a possibilidade de criar uma lógica de controle de leds e diversos dispositivos eletrônicos. Na Figura 4 pode-se observar o circuito de um semáforo sendo controlado via placa Arduino.

## Bibliografia

MESSIAS, A. (2006) "**Porta Paralela: como controlar 8 motores de passo**", <http://www.rogercom.com/pparalela/Lpt8M.html>, Junho.

QUINDERÉ, P. (2009) "**Casa inteligente - Um protótipo de sistema de Automação de baixo custo**", <http://www.ffb.edu.br/sites/default/files/tcc-20082-patrick-romero-frota-quindere.pdf>, Junho.

# Uma Arquitetura para Gestão da Largura de Banda em Redes Domésticas

Luan Teylo G. Lima<sup>1</sup>, Walas Jhony Lopes Moraes<sup>1</sup>, Maxwell Silva Carmo<sup>1</sup>

<sup>1</sup>Instituto de Ciências Exatas e da Terra – Universidade Federal de Mato Grosso (UFMT)

CEP 78600-000 – Barra do Garças – MT – Brasil

{luanteylo, walasjhonyblz}@gmail.com, max@ufmt.br

**Abstract.** *This work proposes an architecture for home network management based on techniques of Software Defined Networking and network virtualization that enable a controlled sharing of bandwidth among users and applications.*

**Resumo.** *Este trabalho propõe uma arquitetura de gerenciamento de redes doméstica que, por meio de técnicas de Software Defined Networking e virtualização de rede, permitirá um compartilhamento controlado de largura de banda entre usuários e aplicações.*

## 1. Introdução

Mesmo com a formidável evolução das tecnologias de rede (*Access Points*, tecnologias em antenas, aprimoramento de padrões, etc.), a baixa largura de banda nas residências brasileiras ainda é um sério problema. Aliado a este, a gestão de uma rede de computadores representa um problema crescente para o usuário doméstico, visto que sua rede está se tornando gradativamente mais complexa. Hoje em um cenário típico, encontram-se conectados em uma mesma rede: uma miríade de dispositivos móveis, impressoras, TVs, consoles de jogos, sistemas de monitoramento (centrais de alarme e câmeras de segurança), além de notebooks e PCs. Para agravar a situação, é comum que a rede seja compartilhada entre várias pessoas, residentes em locais distintos inclusive (compartilhamento de uma rede WiFi entre vizinhos, cenário típico em muitas cidades brasileiras). Esta complexidade tende a aumentar no futuro se for considerado o avanço de tecnologias como *smart grids* e *Internet of things* [Jain et. al 2014]. Por exemplo, um *streaming* de vídeo em alta definição em paralelo com uma aplicação *torrent* poderia trazer efeitos indesejáveis para a experiência do usuário, como atrasos na transmissão e perda na qualidade da imagem.

Este trabalho propõe uma arquitetura para gerenciamento de uma rede doméstica, permitindo compartilhamento controlado de largura de banda para acesso à Internet entre usuários e aplicações. O usuário poderá especificar políticas de uso de alto nível, tais como determinar quando (e quais) aplicações poderão ser executadas na rede (por exemplo, permitir que aplicações *torrent* só sejam executadas a partir das três da manhã ou, como controle de pais, definir que o computador de seu filho não tenha acesso à rede a partir das dez horas da noite), reservar largura de banda às aplicações, e definir quais usuários poderão utilizar a rede (e o quanto da largura de banda poderão consumir e por quanto tempo).

## 2. *Software Defined Networking – SDN*

As redes de computadores atuais podem ser divididas em dois planos de funcionalidade: plano de dados e de controle. O plano de dados é composto pelos dispositivos que transportam os dados (roteadores e *switches*), enquanto o plano de controle representa a inteligência da rede e é responsável por gerenciar os elementos do plano de dados. Nas tradicionais redes IP, os planos de controle e de dados estão acoplados no mesmo dispositivo. Essa arquitetura foi planejada para garantir a resiliência da rede, porém como efeito colateral, tem-se uma arquitetura complexa que dificulta inovações na área. O paradigma SDN supera estas limitações, quebrando a integração entre o plano de controle e o plano de dados. O plano de controle passa a ser logicamente centralizado e pode ser definido como o sistema operacional da rede (NOS, do inglês *Network Operating System*). Isso faz com que os equipamentos que compõem o plano de dados tornem-se simples dispositivos de comutação, isto é, transportadores de pacotes [Kreutz et. al. 2014].

Como o plano de controle é separado do plano de dados, é necessário que exista uma interface entre eles, sendo o *OpenFlow* [McKeown et. al, 2008] atualmente o exemplo mais notável dessa interface, embora não o único. O *OpenFlow* é um protocolo aberto que permite manipular as tabelas de fluxo dos *switches* constituídas tipicamente por: (1) regras para identificar os fluxos através da combinação de vários campos dos pacotes; (2) contadores, utilizados para coletar estatísticas dos fluxos; e (3) um conjunto de ações a serem aplicadas aos fluxos (por exemplo, bloquear todos os pacotes TCP com porta de destino 80, ou encaminhá-los a uma determinada porta do *switch*). A arquitetura aqui proposta fará uso de mecanismos SDN para dar aos usuários o poder de definir políticas de alto nível para controle de suas redes domésticas.

## 3. Compartilhamento de Acesso à Internet

A figura 1 ilustra um cenário típico que pode tirar proveito da arquitetura aqui proposta. Um usuário possui um *link* de alta velocidade de 15Mb com a Internet e para comodidade de conexão dos dispositivos, mantém uma rede WiFi cujo *Access Point* é compatível com o protocolo *OpenFlow*. Ele deseja, eventualmente, compartilhar o acesso à Internet com alguns moradores da redondeza. É imprescindível a este usuário ter meios convenientes de controlar a utilização do *link* de acesso à Internet. Como exemplo, é indesejável que a sincronização de arquivos do *Dropbox* de um vizinho atrapalhe a ligação VoIP feita por outro.

A arquitetura, por meio de técnicas de SDN / virtualização de rede [Chowdhury and Boutaba 2010], permitirá ao proprietário da rede WiFi “fatiar” a largura de banda do *link* de acesso, oferecendo a cada usuário (vizinho) uma capacidade diferenciada. Por exemplo, na figura 1 o convidado B enxerga um *link* de 1Mb, os convidados A e C enxergam 2Mb cada, e o proprietário utiliza a capacidade restante (10Mb). As regras que definem este compartilhamento da rede são atribuídas pelo proprietário a partir do controlador principal. Para cada instância virtualizada será atribuído um software de controle específico, no qual proprietários e convidados poderão criar políticas de uso para suas redes virtuais através de uma interface gráfica simples de utilizar. O usuário poderá definir, por exemplo, que ao executar uma determinada aplicação multimídia o consumo de banda das outras aplicações (*Dropbox*, *Torrent*, etc.) seja temporariamente limitado.

Outro fator a ser considerado é que computadores *desktop* estão cada vez mais ausentes das redes domésticas, sendo substituídos gradativamente por dispositivos móveis. Desta forma, é importante que a arquitetura não dependa de um computador dedicado para executar e armazenar os controladores. Em função disso, serão utilizados serviços como o *Amazon Web Service*, *Windows Azure* ou *Google App Engine*. Alternativamente, usuários responsáveis pela gestão da rede terão a opção de definir dispositivos locais para hospedarem temporariamente os controladores. Uma vez definido tal dispositivo, o controlador será migrado da nuvem para a rede doméstica para processamento local das requisições.

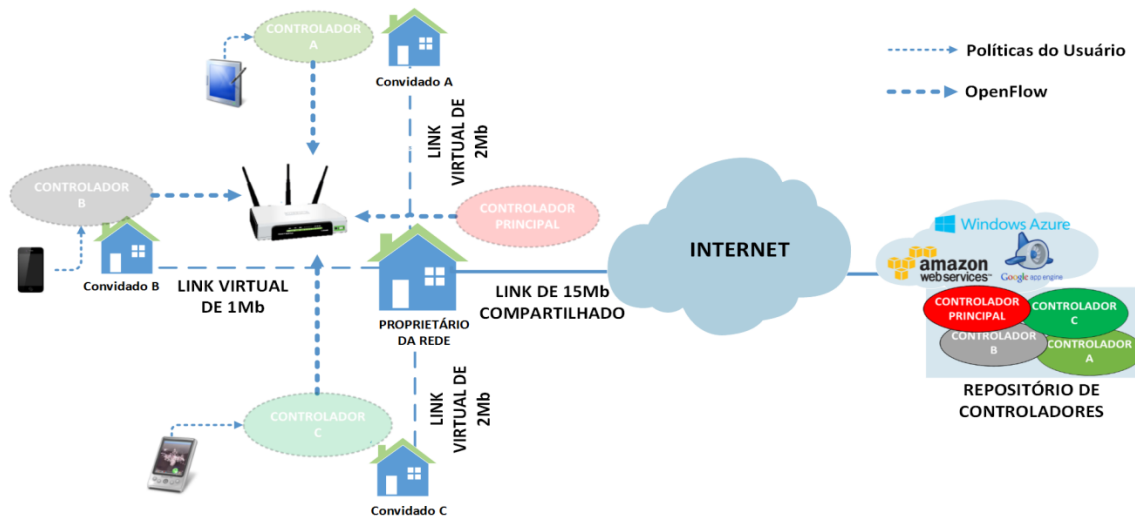


Figura 1. Caso de uso da arquitetura proposta

#### 4. Considerações Finais

Este trabalho encontra-se em fase inicial, na qual os principais detalhes de implementação da arquitetura estão sendo formulados. Atualmente, está sendo desenvolvido um protótipo no qual um *switch* doméstico (TPLink 1043ND) foi transformando em um *switch Openflow* que, quando necessário, contata o controlador localizado nos servidores da *Amazon*. Este protótipo permitirá testes em condições reais, para que seja feita uma estimativa dos custos financeiros e do desempenho geral da arquitetura.

#### 5. Referencias Bibliográficas

- Chowdhury, N. M. M. K. and Boutaba, R. (2010) "A survey of Network Virtualization", In: Computer Networks, p. 862-876.
- Jain, S., Kumar, N.V., Paventhan, A., Chinnaiyan, V.K., Arnachalam, V. and Pradish, M. (2014) "Survey on smart grid technologies- smart metering, IoT and EMS", In: Electrical, Electronics and Computer Science (SCEECS), 2014 IEEE Students' Conference on, p. 1-6.
- Kreutz, D., Ramos, F. M. V., Verissimo, P., Rothenberg, C. E., Azodolmolky, S. and Uhlig, S. (2014) "Software-Defined Networking: A Comprehensive Survey", In: arXiv: 1406.0440, v. 1.0, p. 1-49.
- McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S., Turner, J. (2008) "OpenFlow: Enabling Innovation in Campus Networks", In: ACM SIGCOMM Computer Communication Review, Vol. 38, Issue 2, P. 69-74.

## Inclusão: glossário interativo de Libras

**Lucas C. Borges, Claudia Z. Leite, Fabiano G. Rocha, Wellington M. Cruz,  
Gabriel M. Oliveira, Vinicius S. Ferri**

Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso - *Campus Cáceres*  
Av. dos Ramires, s/n – 78.200-000 – Distrito Industrial – Cáceres – MT – Brasil

{lucas\_campos\_borges, thiagogabrielmalaco, wellington\_morette}@hotmail.com,  
{claudia.leite, fabiano.rocha}@cas.ifmt.edu.br

***Abstract.** The language that the deaf person uses to express themselves and communicate is the Brazilian Sign Language (LIBRAS). To contribute with the process of inclusion of deaf people at the Federal Institute of Mato Grosso - Campus Cáceres, this project consists in developing an interactive glossary of signs referring to the technical area of the courses offered on campus using 3D modeling. It is hoped that the project will collaborate with the inclusion of deaf people, providing better access to technical education.*

***Resumo.** A língua que a pessoa com surdez utiliza para se expressar e comunicar é a Língua Brasileira de Sinais (LIBRAS). Com o objetivo de contribuir com o processo de inclusão de pessoas surdas no Instituto Federal de Mato Grosso - Campus Cáceres, este projeto consiste no desenvolvimento de um glossário interativo de sinais referentes à área técnica dos cursos oferecidos no Campus utilizando modelagem 3D. Espera-se que o projeto venha a colaborar com a inclusão de pessoas surdas, proporcionando um melhor acesso à educação técnica.*

### Introdução

A LIBRAS é a segunda língua oficial do país, oficializada pela Lei nº10.436 de 24 de abril de 2002, usada pela comunidade surda no Brasil. A partir do reconhecimento da LIBRAS, percebemos a necessidade de difusão e formação linguística da Língua de Sinais em todas as instâncias da sociedade, buscando garantir a acessibilidade das pessoas surdas. Os gestos, os sinais e as cores são alguns dos recursos visuais que produzem efeitos e podem ser utilizados na comunicação. Nesse contexto, o uso de tais recursos podem servir de apoio e vir a contribuir para que o sujeito surdo adentre ao “mundo letrado”, como por exemplo em uma instituição escolar.

Para auxiliar no processo de acesso e permanência da pessoa com surdez à instituição escola, existem alguns materiais didáticos e livros adaptados para surdos, contudo, a dificuldade de comunicação é um forte problema quando o aluno surdo adentra na formação profissional. Tal dificuldade deve-se ao fato de que nas disciplinas técnicas são utilizados palavras e termos que não são comuns no cotidiano, e consequentemente, na maioria dos casos, não existe uma representação em Libras.

Neste cenário, surgiu a idéia de desenvolver um glossário interativo de termos técnicos em Libras utilizando computação gráfica em 3D. Com base nos conhecimentos

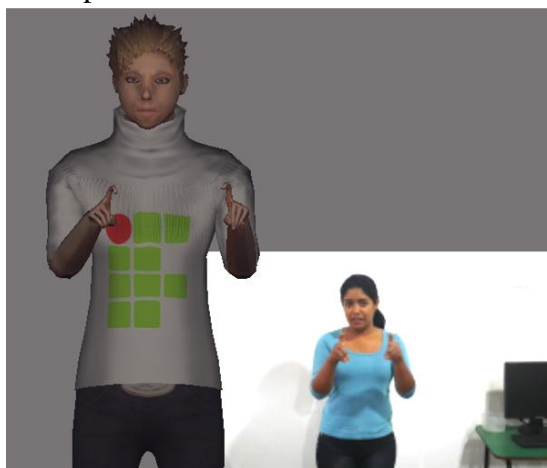
de profissionais de Informática e da Língua de Sinais busca-se mesclar o conteúdo técnico utilizado nos cursos de Informática e a prática de Libras.

## Desenvolvimento

Através do glossário interativo proposto, tanto o aprendiz com surdez quanto o professor e o público em geral podem fazer uso desta ferramenta como apoio às práticas de ensino e aprendizagem. De maneira dinâmica e de fácil acesso, o usuário do glossário pode consultar o significado do termo técnico e sua respectiva representação em Libras por meio da animação 3D desenvolvida.

O projeto está em desenvolvimento no IFMT/*Campus* de Cáceres, no qual foi desenvolvido um avatar em 3D que foi batizado por Stayce. Este avatar tem por função ser o protagonista do glossário devendo reproduzir os termos técnicos que o usuário deseja consultar. A modelagem 3D foi feita utilizando-se o software conhecido por Blender (versão 2.71). O Blender é desenvolvido e mantido pela *Blender Foundation* como software *open source* (código aberto disponível sob a licença GNU GPL) e se destaca no mercado por sua interface simples e a grande quantidade de recursos de modelagem, texturização, iluminação, pós-processamento de vídeo, entre outros.

Para realizar a modelagem 3D, foi necessário criar o mundo em três dimensões, sendo elas: esquerda-direita: linha central X; trás-frente: linha central Y e cima-baixo: linha central Z. A modelagem 3D, que é o resultado da manipulação destas três dimensões, foi criada tendo por referencial dois pontos diferentes (vide Figura 1): frontal e lateral. Este referencial foi obtido pela posição da interprete de libras para reproduzir em 3D o termo técnico correspondente.



**Figura 1: Referencial frontal da interprete de libras e sua correspondência no avatar Stayce.**

Na proposta do glossário, uma vez digitado o termo desejado, o software deve apresentar o seu significado, bem como deve possibilitar ao usuário observar a animação 3D da Stayce reproduzindo o termos correspondente.

Fora efetuado o levantamento dos termos técnicos utilizados na área de Ciência da Computação/Informática tendo, em levantamento preliminar, cerca de 310 termos

técnicos. Em seguida, foi exposto o seu significado e a próxima etapa do projeto será produzir sua equivalente representação na Língua Brasileira de Sinais para então ser modelada em 3D e compor o glossário interativo.

### Resultados Parciais

O projeto encontra-se em andamento e já foi realizado o levantamento dos termos técnicos bem como foi desenvolvido o avatar utilizando a modelagem 3D. A animação 3D do termo técnico em Libras foi desenvolvida no Blender e representada pelo avatar chamado *Stayce*. Na figura 2, pode-se observar um exemplo de tela do glossário interativo no qual tem-se o termo técnico CHAT, seu significado e o usuário pode observar a equivalência do termo técnico em libras na animação 3D.

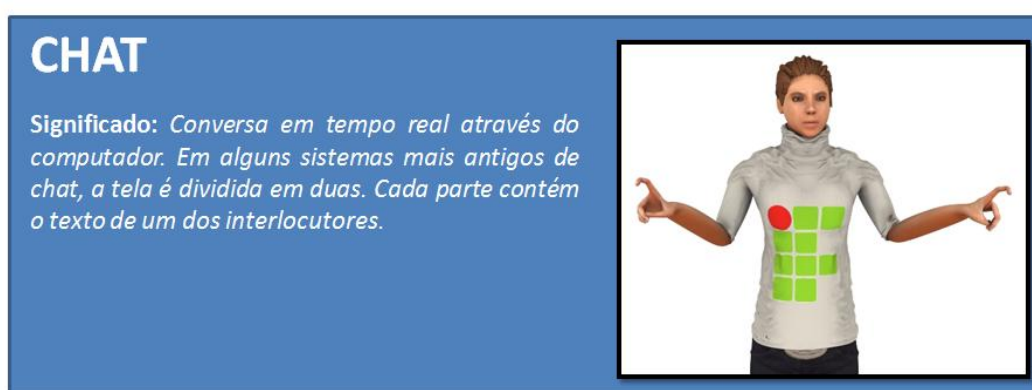


Figura 2: Exemplo de tela do glossário interativo ilustrando o termo técnico CHAT.

### Conclusões Parciais

Acredita-se que com o uso do glossário em desenvolvimento neste projeto, seja possível ajudar, facilitar e difundir os sinais da área técnica no meio acadêmico. Através de recursos visuais (fotos e vídeo) e textos em português, o glossário vem a ser uma ferramenta de apoio pedagógico e de consulta. Espera-se que com o glossário, seja possível oportunizar a comunidade surda um melhor acesso e permanência nos cursos técnicos da Instituição, tendo em vista que teremos uma ferramenta que vem a auxiliar na comunicação, contribuindo diretamente para o resgate da cidadania do sujeito surdo e o seu acesso ao meio técnico-científico.

### Referências

BLENDER. (2014) "**Blender Wiki**", <http://wiki.blender.org>, Agosto.

BOTELHO, P. (2002) "**Linguagem e letramento na educação dos surdos: ideologias e práticas pedagógicas**", Belo Horizonte, Autêntica.

BRASIL. (2002) "**Lei n. 10.436, de 24 de abril de 2002**", Dispõe sobre a Língua Brasileira de Sinais – Libras e dá outras providências, Brasil.

## Proposta da utilização do método *Goal/Question/Metric* para avaliação de acessibilidade em uma Biblioteca Virtual

Lyncon Vieira Lima<sup>1</sup>, Abraão Lincon da Silva Ramalho<sup>1</sup>, Bruno Silva Macário<sup>1</sup>, Leticia Cristina Amorim de Assis<sup>1</sup>, Roan Siviero Dotto<sup>1</sup>, Juliana Fonseca Antunes<sup>1</sup>

<sup>1</sup>Departamento de Informática – Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) - Campus Cuiabá - Octayde Jorge de Silva  
Caixa Postal 78.005-200 – Rua Profª Zulmira Canavarros, N°95 – Cuiabá – MT

{lynconvieira,abraao.lincon2310,br.s.m.10,leticia.aassis,roansdotto}@gmail.com, juliana.antunes@cba.ifmt.edu.br

**Resumo.** *Cada vez mais tem-se utilizado a Internet para buscar informações, porém muitas vezes as informações não estão acessíveis para usuários com deficiência visual pelo fato dos sites não estarem preparados para os mesmos. Este trabalho tem como objetivo propor a utilização da técnica Goal/Question/Metric para avaliar a aderência da acessibilidade em uma biblioteca virtual baseando-se nas diretrizes de acessibilidade da WCAG. Acredita-se que definindo objetivos, questões e métricas pode-se realizar uma avaliação com maior ponderação a fim de aperfeiçoar o acesso à informação para esses usuários especiais.*

**Abstract.** *Increasingly it has been used the Internet to search for information, but often the information is not accessible to visually impaired users because the sites are not ready for them. This paper aims to propose the use of the technique Goal/Question/Metric to assess the adherence of accessibility in a virtual library building on the accessibility guidelines of WCAG. It is believed that setting goals, questions and metrics can perform an assessment with greater weight in order to improve access to information for these special users.*

### 1.Introdução

A acessibilidade vem crescendo acentuadamente no Brasil; Muitos restaurantes, bancos, cinemas e outras instituições públicas estão disponibilizando esse acesso a quem necessita (Souza, F.; Osorio, M.; Andrade, E.; 2008). Hoje, com o rápido avanço da tecnologia, já não é mais suficiente apenas a acessibilidade física para os deficientes, mas também é de grande importância existir a acessibilidade *web*. Acessibilidade na *web* significa que cidadãos com qualquer deficiência possam utilizar, entender, contribuir, interagir e navegar na Internet sem qualquer empecilho (Shawn, 2005).

Quando se trata de acessibilidade *web*, os mais prejudicados são as pessoas com limitações visuais pois a *web* é altamente visual, contendo muitos textos, imagens, vídeos, etc. As necessidades que uma pessoa com deficiência visual tem ao navegar pelos *sites* são semelhantes à de uma pessoa sem essa deficiência: se atualizar com as últimas notícias, se entreter através de jogos, realizar operações bancárias, fazer compras, etc. A acessibilidade é um

dos fatores que compõem a qualidade de um site, por isso a preocupação de diversas instituições e governos em relação a este assunto (Lima, S., Lima, F., Oliveira, K., 2007).

É necessário uma atenção maior na hora de projetar um *site* acessível para usuários desse tipo, pois nota-se que os problemas mais sérios relacionados à acessibilidade são pertinentes a deficientes visuais. É necessário que o usuário, seja ele deficiente ou não, entenda o que está sendo escrito ou mostrado no site para ocorrer à interação com o sistema. Conforme NIELSEN e LORANGER (2007), o projetista deve ter como regra básica o conhecimento do usuário antes de qualquer outra coisa.

“Quando tornamos o nosso *site* acessível, além de atingirmos os usuários da Internet que não podiam acessar devido às barreiras encontradas, também estamos criando condições para que novas pessoas se animem a usar a Internet; ou seja, estamos ampliando o nosso mercado. (...)” (Spelta, 2008).

Considerando que o acesso à informação na Internet deve ser igual para qualquer pessoa, independente de possuir a capacidade visual ou não, este trabalho tem como objetivo avaliar a aderência de uma Biblioteca Virtual aos níveis de acessibilidade aceitáveis, utilizando a técnica *Goal/Question/Metric*. A ideia básica do *GQM* é derivar métricas de software a partir de perguntas e objetivos como parte de iniciativa de melhoria geral nos processos de desenvolvimento de software. O foco principal da prática é definido em métricas direcionadas a metas, mas também aborda a coleta de dados, análise e interpretação e armazenamento das experiências para uso em futuras iniciativas.

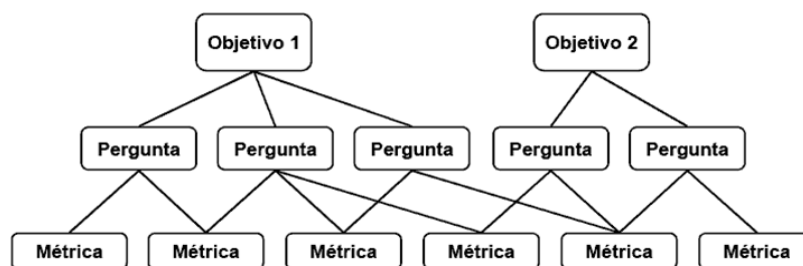
## 2. Objetivo

Este trabalho tem por objetivo propôr a utilização da técnica *Goal/Question/Metric* definindo objetivos, questões e métricas para avaliar a aderência da acessibilidade de uma Biblioteca Virtual baseando-se nas diretrizes de acessibilidade da *WCAG (Web Content Accessibility Guidelines)* afim de aperfeiçoar qualitativamente o acesso às informações para os/esses usuários com deficiência visual.

## 3. Metodologia

Métricas podem ser utilizadas para caracterizar e aperfeiçoar processos, produtos e recursos. Usadas para medir, por exemplo, a qualidade característica de um determinado produto de software, a eficácia de um determinado processo ou a quantidade de recursos gastos em um produto com um tamanho específico. Um fator empírico é adicionado à pesquisa sobre engenharia de software por meio de métricas: novos métodos, técnicas e ferramentas são medidas para compreender como eles são bons, quais problemas resolvem, quais não resolve, e as necessidades básicas para aplicá-las. Em outras palavras: métricas fornecem um excelente mecanismo para saber o que funciona e o que não funciona.

A ideia básica de *GQM* é derivar métricas de software a partir de perguntas e objetivos. Usando *GQM* um determinado objetivo é definido, este objetivo é refinado em perguntas, e posteriormente, as métricas que são definidas devem fornecer as informações para responder estas questões. Ao respondê-las, os dados de medição podem ser analisados para identificar se os objetivos são atingidos. Assim, usando *GQM*, métricas são definidas a partir de uma perspectiva *top-down*, analisadas e interpretadas com *bottom-up*, conforme mostra a FIGURA 1.

FIGURA 1: Estrutura da técnica *Goal/Question/Metric*

Os dados reais podem então ser usados para comparar com as hipóteses iniciais e identificar causas para as falhas e aprender relações causais.

Com base na técnica descrita acima, propomos definir questões e métricas para a avaliação da acessibilidade na biblioteca virtual, sob o ponto de vista de um usuário com deficiência visual. Relatar questões como: sequência de tabulação da página em uma escala qualitativa, por exemplo, criando-se métricas de nível de satisfação em relação a sequência. Visando assim minimizar as barreiras impostas aos deficientes visuais melhorando o acesso a informação.

#### 4. Conclusão

Embora o benefício direto do *GQM* seja estabelecer métricas significativas, quando o *GQM* é aplicado no contexto da melhoria do processo sistemático, é especialmente útil para avaliar e certificar melhorias nas atividades (Silva, C.; Campos, D.; Nery, P.; Moura, D.; 2009).

Como perspectivas futuras, espera-se que com a aplicação desta técnica e o auxílio das diretrizes de acessibilidade da *WCAG* seja possível a realização de uma avaliação com cautela dos obstáculos que serão encontrados na biblioteca virtual, promovendo assim um aperfeiçoamento para evitar que seja repetido os mesmos erros na correção, tornando-a acessíveis a pessoas com deficiências visuais.

#### 5. Referências

SPELTA; LÊDA. Acessibilidade web: 7 mitos e um equívoco. Disponível em: <[http://acessodigital.net/art\\_acessibilidade-web-7-mitos-e-um-equivoco.html](http://acessodigital.net/art_acessibilidade-web-7-mitos-e-um-equivoco.html)>. Acesso em: 20 ago 2014.

LIMA, S.; LIMA, F.; OLIVEIRA, K., Avaliação da Acessibilidade de Sítios Web por meio de Métricas de Software. VI Simpósio Brasileiro de Qualidade de Software, 2007.

SHAWN L.H, and participants of the education and Outreach Working Group (EOWG) (2005) "Introduction to Web Accessibility". Disponível em: <<http://www.w3.org/WAI/intro/accessibility.php>>. Acesso em: 24 ago 2014.

NIELSEN, J.; LORANGER, H. – Usabilidade na Web – Projetando Websites com qualidades. Rio de Janeiro, RJ. Elsevier Editora Ltda. 3ª reimpressão, 2007.

SOUZA, F.; OSÓRIO, M.; ANDRADE, E. – Proposta de avaliação de acessibilidade em sites por deficientes visuais. 2008.

SILVA, C.; CAMPOS, D.; NERY, P.; MOURA, D. – *GQM Goal – Question – Metric*. 2009

# Acesso Compartilhado à Internet através de Redes WiFi

Walas Jhony Lopes Moraes<sup>1</sup>, Luan Teylo Gouveia Lima<sup>1</sup>, Maxwell Silva Carmo<sup>1</sup>

<sup>1</sup>Universidade Federal de Mato Grosso

walasjhony@outlook.com, luanteylo@gmail.com, max@ufmt.br

**Abstract.** *Considering the large number of WiFi networks existing in the urban centers currently, we propose an architecture that allows domestic users to use those networks for access to the Internet, wherever each user shares some network resources. Our goal can be achieved using Software-Defined Networking techniques. It provides us ways to manage dynamic and rational those resources.*

**Resumo.** *Levando em conta o grande número de redes WiFi presentes nos centros urbanos atualmente, propomos uma arquitetura que permite aos usuários em geral fazer uso destas redes para acesso à Internet, tendo como norte o compartilhamento voluntário dos recursos de rede. Tal objetivo é alcançado por intermédio de técnicas de Software-Defined Networking (SDN), que provê meios para o gerenciamento dinâmico e racional de tais recursos.*

## 1. Introdução

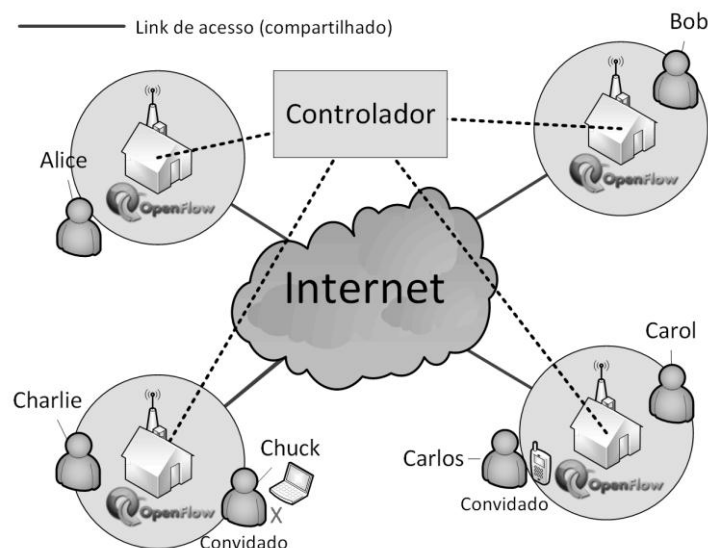
É comum, nos dias de hoje, residências possuírem uma rede doméstica WiFi com acesso à Internet. Com a tecnologia vigente, essas redes podem ser acessadas somente dentro de um raio bastante limitado. Caso um usuário queira acessar a Internet a partir de qualquer lugar, é necessário adquirir outro plano de dados (EDGE, 4G, etc.), que em geral apresenta desempenho aquém do desejado.

A alta densidade de redes WiFi em um centro urbano contrasta enormemente com sua falta de disponibilidade: elas estão presentes, mas inacessíveis ao usuário em geral. Neste trabalho propomos uma arquitetura que procura fazer uso inteligente dos recursos destas redes para fornecer acesso à Internet a uma ampla gama de usuários. A ideia que dá sustentação à proposta não é nova: O proprietário de uma rede WiFi doa parte de seus recursos (especificamente, largura de banda para acesso à Internet) a outros usuários e em troca ele pode utilizar recursos de outra rede como convidado, assim que estiver fisicamente dentro de seu alcance. Ou seja, os recursos são providos pela comunidade usando a infraestrutura de rede já existente. Quanto maior o número de usuários a utilizar a arquitetura, maior a disponibilidade de acesso (quantidade de redes disponíveis). Projetos como FON [Asheralieva, Erke e Kilkki 2009] já vêm seguindo este modelo com sucesso considerável. O que diferencia o trabalho aqui apresentado dos demais é a utilização da tecnologia de *Software Defined Networking* (SDN) [Kreutz et al. 2014] para gerir de forma dinâmica e racional os recursos das redes WiFi. Por exemplo, um usuário convidado poderá reservar temporariamente recursos da rede (especificamente, largura de banda da rede de acesso) para fazer uma chamada por vídeo sem sofrer a interferência do tráfego proveniente de outros usuários, que também estão acessando a Internet através da mesma rede WiFi. Assim que concluir a chamada, a largura de banda reservada volta ao *pool* de recursos disponíveis.

## 2. A arquitetura

A Figura 1 ilustra o cenário de operação da arquitetura. O proprietário de uma rede WiFi com conexão à Internet delega uma fatia da largura de banda de seu *link* de acesso ao controle da arquitetura. Quando outro usuário, chamado aqui de convidado, se encontra ao alcance desta rede, a arquitetura lhe permite utilizar a fatia que está sob sua tutela. A arquitetura centraliza informações sobre todas as redes WiFi que podem ser compartilhadas. A gestão dos recursos delegados a ela também ocorre de forma logicamente centralizada, embora deverão ser empregados componentes redundantes (confiabilidade) e dispostos hierarquicamente (escalabilidade).

O proprietário da rede tem total autonomia para delegar ao controlador a quantidade de recurso que deseja compartilhar e por quanto tempo. É fato que o sucesso da proposta vai depender do incentivo que os usuários terão para dispor de seus recursos. Quanto maior a largura de banda e o número de redes WiFi que a arquitetura gerencia, maior o número de usuários convidados atendidos e a área geográfica abrangida. A princípio, pode-se conceber uma política “olho por olho” inspirada no protocolo de compartilhamento de arquivos BitTorrent. Usuários que disponibilizam mais recursos têm direito de consumir mais recursos. Esta política precisa ser refinada em função das particularidades dos recursos aqui compartilhados. Por exemplo, uma rede WiFi localizada no centro da cidade será mais requisitada do que uma situada em uma zona pouco habitada de um subúrbio. Embora ambas possam disponibilizar a mesma quantidade de recurso, a segunda pode não ser requisitada por nenhum convidado por longos períodos. No momento de contabilizar o quanto um convidado tem direito, fatores como estes precisam ser levados em conta. Além disso, é preciso ser capaz de identificar usuários que não disponibilizam recursos (não fazem parte da comunidade) ou cuja cota já chegou nos limites calculados pela arquitetura. Para ambos os casos tentativas de conexão a qualquer rede devem resultar em bloqueio de acesso.



**Figura 1. Cenário de operação da arquitetura**

Utilizando-se de técnicas de SDN e de virtualização de redes [Sherwood et al. 2009], a arquitetura irá garantir o isolamento entre os recursos sob sua responsabilidade e os do usuário. Por exemplo, caso o proprietário resolva delegar 2Mb dos 10Mb disponíveis em seu *link* de acesso, a proporção será mantida.

Com o auxílio de um componente que executa no dispositivo móvel do convidado, decisões como a escolha da melhor rede a se conectar entre as disponíveis no momento serão tratadas pela arquitetura, exigindo o mínimo possível de interferência do usuário. Uma vez que a arquitetura contabiliza informações de compartilhamento de todas as redes WiFi, ela pode sugerir uma rede em detrimento de outra de acordo com fatores como proximidade do AP e condições de uso da rede no momento (se outro convidado já está conectado à rede A, talvez seja melhor utilizar a rede B).

Para oferecer os serviços propostos acima é necessário ao *Access Point* ser compatível com o protocolo OpenFlow [McKeown et al. 2008]. Tais equipamentos são encontrados com relativa facilidade no mercado hoje em dia. Por exemplo, a implementação do protocolo no sistema openWRT [Palazzi, Brunati e Rocchetti 2010] permite sua utilização em uma série de APs.

As interfaces do usuário devem permitir que este, sem conhecimento avançado, possa configurar seu AP (quando e quanto da sua largura de banda poderá ser usado por convidados) e acessar uma rede WiFi como convidado.

### 3. Considerações finais

Na altura da escrita deste resumo aspectos arquiteturais importantes da proposta (que se encontrar em estágio de modelagem) ainda estão sendo analisados. Soluções eficientes para os problemas de escalabilidade (milhares de APs localizados possivelmente em cidades diferentes estarão sob o controle da arquitetura), confiabilidade (se o controlador deixar de funcionar abruptamente, é necessário substituí-lo sem grandes prejuízos aos usuários) e distribuição dos recursos com equidade precisam ser mais bem definidas. Após esta etapa inicial de *design*, um protótipo será desenvolvido utilizando o ambiente Mininet [Mininet.org] para validação da arquitetura.

### Referências

- Asheralieva, A., Erke, T. J., e Kilkki, K. (2009). "Traffic characterization and service performance in FON network", Em: *International conference on future information networks*, p. 285-291.
- Kreutz, D., Ramos, F. M. V., Verissimo, P., Rothenberg, C. E., Azodolmolky, S. e Uhlig, S. (2014) "Software-Defined Networking: A Comprehensive Survey", Em: *arXiv*: 1406.0440, v. 1.0, p. 1-49.
- McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S. e Turner J. (2008) "OpenFlow: Enabling Innovation in Campus Network", Em: *ACM SIGCOMM Computer Communication Review*, Vol. 38, Issue 2, p. 69-74.
- Mininet.org. Disponível em: <<http://mininet.org/overview>>. Acesso: 26 ago. 2014.
- Palazzi, E.C., Brunati, M. e Rocchetti, M. (2010) "An OpenWrt solution for future wireless homes", Em: *IEEE International Conference on Multimedia and Expo*, p. 1701-1706.
- Sherwood, R., Gibb, G., Kiong, K., Appenzeller, G., Casado, M., McKeown, N. e Parulkar, G. (2009) "FlowVisor: A Network Virtualization Layer", <http://openflowswitch.org/downloads/technicalreports/openflow-tr-2009-1-flowvisor.pdf>

## Usabilidade nos Portais dos Institutos Federais do Centro-Oeste

Beatriz de P. Moraes<sup>1</sup>, Thatielle M. Pereira<sup>1</sup>, Lorena L. Oliveira<sup>1</sup>,  
Carine R. da Costa<sup>1</sup>

<sup>1</sup>Instituto Federal de Educação, Ciência e Tecnologia de Mato Grosso (IFMT) – Campus Barra do Garças  
78600-000 – Barra do Garças – MT – Brasil

{bia\_moraes182,fabiula1711}@live.com, loreloliveira@hotmail.com,  
carine.rodrigues@bag.ifmt.edu.br

**Abstract.** *The Web page of a Federal Institute of Education is a representation of your identity or biography. It should contain all possible information from the Federal Institute in an organized fashion in order to simplify and facilitate access to them. The Web page structure and the information presentation must be defined using the concept of usability. The research presented below has as main objective the study of the concepts of usability of different Federal Institutes Web pages, verify the possibility of rapid and practical methods of usability evaluation and identify usability issues in these portals.*

**Resumo.** *Um portal de Instituto Federal é a representação da identidade ou biografia do mesmo, onde nele deve conter todas as informações possíveis de maneira organizada, simplificando e facilitando o acesso. A estrutura dos portais e apresentação das informações devem ser definidas a partir do conceito de usabilidade. A pesquisa apresentada a seguir tem como principais objetivos o estudo dos conceitos de usabilidade de portais de diferentes Institutos Federais, verificar a possibilidade de métodos rápidos e práticos de avaliação de usabilidade e identificar os problemas de usabilidade dos respectivos portais.*

### 1. Introdução

A usabilidade é a maneira mais prática e simples possível de interação do internauta com o site (CARVALHO, 2013). Esta é representada por meio de uma interface bem planejada, onde as regras básicas de navegabilidade, acessibilidade e usabilidade deverão estar implementadas.

Caso o site implemente as regras de usabilidade, estará apto a atender seu público alvo, pois ele terá disponível as informações necessárias que são procuradas pelo usuário, além de ter maior facilidade e disponibilidade de acesso sem transtornos. Isso garantirá satisfação ao usuário, que terá melhor rendimento em sua pesquisa, enfim, no alcance de seus objetivos. Na Figura 1 e na Figura 2, que apresentam os portais do

IFMT e IFB, respectivamente, é possível ver as diferenças na estruturação e organização da informação.

## 2. Metodologia

Foi realizada uma pesquisa bibliográfica sobre conceitos de usabilidade para páginas web e métodos e/ou técnicas de avaliação de usabilidade. Foram escolhidos os portais dos institutos federais da região Centro-Oeste, que são: IFMT, IFMS, IFB, IFG e IFGoiano. A escolha foi devido à similaridades que podem ser encontradas na região, que impactam na formação do perfil e identidade de cada instituto.

## 3. Resultados Parciais

Foram pesquisados métodos para avaliação de usabilidade de sites. Os métodos mais utilizados são: analítico, experimental e de pesquisa de opinião (KRONE, 2013). Uma das técnicas analíticas mais utilizadas é a inspeção por meio de listas de verificação (conferência), onde não há a necessidade de especialistas e interação com usuário. O resultado da aplicação dessa técnica depende da qualidade da lista, que deve ser bem definida e pertinente para as avaliações (CIBYS et. al, 2010).

O método escolhido foi o analítico. Como lista de verificação, foi encontrado o ErgoList, desenvolvido pela Universidade Federal de Santa Catarina, que contém um questionário online, com vários critérios de análise. Ao utilizar o checklist, é gerado um lado final sobre o portal ou site (CIBYS, 2011).



Figura 1. Portal do IFMT. Disponível em: <http://www.ifmt.edu.br/>



Figura 2. Portal do IFB. Disponível em: <http://www.ifb.edu.br/>

#### 4. Conclusões

Em relação ao Ergolist, a necessidade de estabelecer atributos e itens para aplicação do *checklist* é necessária. Devem ser abordados inicialmente alguns características de eficácia, eficiência ou satisfação. Como passos futuros, serão aplicados os testes de usabilidade do Ergolist, pois a importância da comparação de *websites* pode contribuir não somente com avaliação de ergonomia dos portais e o levantamento informações a cerca da arquitetura da informação contida neles, mas também a reestruturação da identidade do portal, sempre valorizando suas especificidades, como região, corpo docente, demandas regionais, etc.

#### Referências Bibliográficas

- CARVALHO, Henrique. “18 Problemas de Usabilidade que deixam seus leitores furiosos (E como consertá-los)”. Disponível em: <http://viverdeblog.com/18-problemas-usabilidade>. Acesso em: 09 abr 2014.
- CYBYS, Walter. ErgoList. 2011. Disponível em: <http://www.labiutil.inf.ufsc.br/ergolist/>. Acesso em: 29 set 2014.
- CIBYS, Walter. BETIOL, Adriana H. FAUST, Richard. “Ergonomia e Usabilidade: Conhecimentos, Métodos e Aplicações”. 2 ed. Rev. e Amp. São Paulo: Novatec Editora, 2010.
- KRONE, Caroline. “Validação de Heurísticas de Usabilidade para Celulares Touchscreen”. Grupo de Qualidade de Software – UFSC. Florianópolis, 2013.

## **ELiS Mobile: trazendo a comunicação escrita dos surdos para os dispositivos móveis**

**Pedro Francisco de Sousa Neto<sup>1</sup>, Marcelo Ricardo Quinta<sup>1</sup>, Mariângela Estelita Barros<sup>2</sup>**

<sup>1</sup>Instituto de Informática – Universidade Federal de Goiás (UFG)  
Avenida Esperança, Vila Itatiaia – Câmpus Samambaia – Goiânia – Goiás – Brasil

<sup>2</sup>Faculdade de Letras – Universidade Federal de Goiás (UFG)  
Avenida Esperança, Vila Itatiaia – Câmpus Samambaia – Goiânia – Goiás – Brasil

{pedrokra,marceloricardoquinta}@gmail.com, estelitabrasil@gmail.com

***Abstract.** Mobility brings to users more opportunities to promote interaction and may experience improvement in software products. This improvement can also be offered to people with disabilities through accessible apps or those ones whose scope are produced only to assist the disabled in special activities. Thinking about it, this study presents a ELiS project, whose goal is to facilitate the written deaf communication in a language closer to the signs used in LIBRAS (Brazilian Sign Language).*

***Resumo.** A mobilidade traz aos usuários mais oportunidades de interação e pode promover melhoria da experiência em produtos de software. Tal melhora também pode ser oferecida a pessoas com deficiência através de aplicativos com características acessíveis ou que sejam produzidas para auxiliar em atividades especiais aos deficientes. Pensando nisso, este trabalho apresenta o projeto ELiS, cujo objetivo é facilitar a comunicação escrita dos surdos através de uma linguagem mais próxima aos sinais utilizados em LIBRAS (Língua Brasileira de Sinais).*

### **1. Introdução**

A Língua Brasileira de Sinais (Libras) ganhou maior visibilidade no Brasil depois que a Lei nº 10.436/2002 foi sancionada, reconhecendo-a como língua da comunidade surda brasileira. Tal lei reconhece também “outros recursos de expressão a ela associados” (Art.1º).

Um dos meios de expressão intimamente relacionado a Libras é sua escrita, no entanto, esta tem sido ainda negligenciada por cursos livres que ensinam a Libras ou escolas regulares que atendem a crianças surdas. Uma das linguagens de representações de sinais em escrita é a ELiS, criada no ano de 1998 pela linguista Mariângela Estelita Barros (Barros, 2008). Essa escrita é de base alfabética, cada letra/símbolo é denominado visografema, elementos visuais formadores dos sinais.

Entende-se que a falta de suporte existente se dê em parte pela falta de preparo de professores para ensinar sua escrita, em parte por falta de material didático ou de consulta disponível para apoiar essa prática. É necessário oferecer meios e ferramentas para acesso e ensino ao conteúdo, por exemplo, um dicionário com tradução, como uma ferramenta para facilitar a comunicação.

Uma forma de sanar este tipo de necessidade é através da distribuição de aplicativos móveis, já uma realidade em nosso cotidiano, seja para acessar extensões de aplicativos já existentes na web ou em computadores de mesa, conhecer novas funcionalidades que só a mobilidade pode nos trazer (i.e. GPS) ou acessar conteúdo.

Logo, este trabalho apresenta a solução ELiS Mobile, um tradutor ELiS-Português para acesso às palavras do dicionário.

## **2. O desenvolvimento do *ELiS Mobile***

O ELiS Mobile é um aplicativo Android para dispositivos móveis. Ele contém a funcionalidade de tradução de palavras e tem como entrada termos em português ou em ELiS. Para que o aplicativo seja executado, é necessário que o dispositivo esteja com o sistema operacional Android na versão 2.3.3 ou superior e de acesso à internet, para que as consultas à base de dados possam ser realizadas através de um servidor na nuvem.

A metodologia deste trabalho segue o padrão do desenvolvimento tradicional de sistemas: definição de requisitos e arquitetura, seguidos da codificação e design.

Estruturação dos visografemas: Sendo a ELiS a escrita das línguas de sinais, os visografemas (elementos básicos formadores de sinais) têm uma representação especial, com fonte personalizada e que pode ter estilo sublinhado ou sobrescrito. Para que isso fosse possível no aplicativo, a fonte tipográfica foi transformada em web-font, fazendo com que o texto pudesse ser mostrado corretamente em HTML, compatível com a web e campos de texto no Android.

Mecanismos para a consulta via serviços web por palavras e símbolos - Assim como vários sistemas existentes na academia e mercado, serviços REST (Fielding, 2000) foram definidos para permitir o acesso a diferentes tipos de clientes sem grande acoplamento de tecnologias, o que permite ao sistema crescer no futuro. O serviço mais utilizado é o de busca de palavra, onde o usuário escolhe a forma de escrita e visualização do resultado, tendo a resposta do servidor no formato JSON (Javascript Object Notation).

Criação de interfaces: A criação de interfaces compreendeu a criação das interfaces gráficas do aplicativo no Android, utilizando bibliotecas nativas do Google. Um ponto importante nesta etapa foi a concepção do teclado ELiS, que contém mais de 90 caracteres especiais, fora as teclas de formatação. Os símbolos foram divididos em grupos contextuais relacionados aos sinais. A escolha dos caracteres pode ser feita utilizando gestos ou toques na tela.

Desenvolvimento de código: Criação de artefatos de códigos que implementam os requisitos, dadas as restrições arquiteturais e visuais.

Testes de implantação: Testes realizados dentro da comunidade de surdos.

## **3. Conclusão**

O desenvolvimento do aplicativo por si só já é um resultado válido. A partir deste, os surdos poderão entender e praticar mais a escrita desta língua, o que não era possível anteriormente em nenhum aplicativo da Apple Store (Apple, 2014) e Play Store (Google, 2014) até Agosto de 2014. Além disso, a construção de um layout padrão para entrada de visografemas em um dispositivo móvel também é um resultado interessante,

pois validou a possibilidade de escrita em um layout de teclado diferente do convencional - QWERTY.

Como a ELiS concebeu seus visografemas relacionados aos sinais utilizados em LIBRAS, um resultado indireto esperado é o conhecimento dos visografemas mais utilizados nas palavras do cotidiano. Isso só será possível futuramente, quando uma massa maior de pesquisas de palavras for realizada.

Após a definição de requisitos do aplicativo, novas ideias já surgiram e ainda serão implementadas, como pesquisa lexicográfica. Esta permitirá que profissionais da linguística também possam explorar a lista de palavras e seus metadados. Outra contribuição que pode acontecer é a adição de vídeos dos sinais em LIBRAS à resposta do servidor.

Um módulo de jogo ainda está sendo desenvolvido. Ele é visto como uma ferramenta descontraída de ensino da língua. Acredita-se que o conhecido jogo “forca” seja o mais indicado, visto sua simplicidade e capacidade lúdica de ensinar a relação visografema-palavra.

## Referências

APPLE “App Store”. Disponível: <https://itunes.apple.com/br/genre/ios>, Agosto/2014.

BARROS, M. E. ELiS – Escrita das Línguas de Sinais: proposta teórica e verificação prática. 192f. Tese (Doutorado em Linguística) – Centro de Comunicação e Expressão, Universidade Federal de Santa Catarina, Florianópolis, 2008.

FIELDING, Roy Thomas. Architectural Styles and the Design of Network-based Software Architectures. Doctoral dissertation, University of California, Irvine, 2000.

GOOGLE “Google Play”. Disponível: <https://play.google.com/store>, Agosto/2014.