

## SISTEMAS DE RECOMENDAÇÃO BASEADOS EM WEB USAGE MINING

Fábio A. Procópio de Paiva<sup>1</sup> e José A. F. Costa<sup>2</sup><sup>1</sup> Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte – Campus Zona Norte<sup>2</sup> Universidade Federal do Rio Grande do Norte – Departamento de Engenharia Elétrica

fabio.procopio@ifrn.edu.br – jafcosta@gmail.com

Artigo submetido em maio/2012 e aceito em junho/2012

## RESUMO

Diariamente, uma quantidade enorme de páginas é publicada na Web e, conseqüentemente, a dificuldade encontrada por usuários para localizá-las é cada vez maior. O desafio é, então, facilitar essa localização e, sobretudo, selecionar apenas informações que são, de fato, relevantes. Portanto, é evidente a necessidade dos usuários em utilizar mecanismos personalizados que apresentem informações adequadas, e ofereçam recomendações coerentes. Na tentativa de

oferecer isso aos inúmeros usuários web, nos últimos anos, várias pesquisas estão sendo realizadas nas áreas da *Web Usage Mining* (WUM). Neste artigo, fizemos uma pesquisa de algumas abordagens de sistemas de recomendação web que estão disponíveis na literatura e um dos nossos objetivos é contribuir com trabalhos de outros pesquisadores dessa área.

**PALAVRAS-CHAVE:** *Web Usage Mining*, interesse de usuário, sistema de recomendação, comportamento de navegação

## RECOMMENDER SYSTEMS BASED ON WEB USAGE MINING

## ABSTRACT

A huge amount of pages is published on the Web every day. Consequently, users find it ever more difficult to localize those pages. The challenge is to facilitate this localization and overall to select only information that is actually relevant. Therefore, there is an evident need for users to utilize personalized mechanisms that present adequate information and offer coherent recommendations. In the attempt of

providing such to the countless web users, in the last years, several researches are being conducted in the areas of *Web Usage Mining* (WUM). In this paper, we have conducted a research on several approaches to web systems of recommendation available in the literature and one of our goals is to contribute to the work of other researchers in the area.

**KEY-WORDS:** *Web Usage Mining*, user interest, recommendation system, navigational behavior

---

## SISTEMAS DE RECOMENDAÇÃO BASEADOS EM WEB USAGE MINING

### INTRODUÇÃO

O conteúdo disponível na WWW está armazenado em diversos formatos, abrange distintas áreas do conhecimento e contempla usuários com vários perfis de interesse. Contudo, o desafio é tornar o acesso às informações rápido e, sobretudo, torná-las relevantes aos interesses dos usuários. Assim, a *World Wide Web* configura-se como um ambiente propício para aplicação das técnicas de Mineração de Dados, conhecida como *Web Mining*. Relacionada com vários outros domínios do conhecimento – dentre eles, Processamento de Linguagem Natural, Banco de Dados, Recuperação da Informação e Inteligência Artificial –, a *Web Mining* consiste na extração de padrões considerados interessantes de navegação, bem como na identificação de informações implícitas nos recursos WWW. Diversas pesquisas vêm sendo realizadas nessa área e, por isso, vem se tornando bastante promissora e uma das mais importantes da Ciência da Computação e da Informação (Pamnani e Crawan, 2010).

Existem várias abordagens que podem ser utilizadas para identificar o comportamento de navegação dos usuários web. Diversas propostas encontradas na literatura utilizam técnicas de Inteligência Artificial como Colônia de Formigas, Algoritmo *Naïve Bayseano*, Mapas de Kohonen e outras. Ling *et. al.* (2007) construíram um modelo baseado em colônias de formigas para identificar os padrões de navegação dos usuários e, conseqüentemente, mapear os seus interesses. Na proposta, os autores consideram usuários web como “formigas artificiais” e usam o comportamento delas como metáfora para guiar a escolha feita por eles. Etmnani *et. al.* (2009) utilizaram Mapas de Kohonen para apresentar uma maneira de detectar o comportamento navegacional dos usuários web.

Com o objetivo de auxiliar pesquisadores que estudam a *Web Usage Mining* e suas aplicações, este artigo apresenta algumas abordagens propostas para sistemas de recomendação baseados em técnicas de *Web Usage Mining*. O trabalho está organizado da seguinte forma: nas seções WEB MINING e WEB USAGE MINING são apresentados os principais conceitos dessas áreas. Na seção seguinte, sobre trabalhos desenvolvidos, elencamos algumas propostas publicadas na área de sistemas de recomendação. Por fim, na seção CONCLUSÕES, fazemos nossas considerações finais.

### WEB MINING

A *Web Mining*, ou Mineração na Web, aplica as técnicas de Mineração de Dados ao conteúdo, à estrutura e à utilização de recursos disponíveis na Web. Podendo ser aplicada a dados estruturados ou semi-estruturados e, ainda, a dados no formato de texto, a Mineração Web propicia que o conteúdo seja compreensível tanto por seres humanos quanto por máquinas (Stumme *et. al.*, 2006). A *Web Mining* é classificada em 03 (três) categorias (Pani, 2011): *Web Content Mining*, *Web Structure Mining* e *Web Usage Mining*.



Figura 1 – Principais aplicações da *Web Usage Mining*

A **Web Content Mining** tem o objetivo de descobrir as informações e os conhecimentos úteis, inseridos nos conteúdos das páginas web. A categoria **Web Structure Mining** estuda como os *hiperlinks* estão estruturados na Web. E, por fim, a **Web Usage Mining** consiste na extração de padrões de utilização de informações web.

Atualmente, a **Web Usage Mining** é uma das áreas de interesse de muitos pesquisadores e isso é o resultado de algumas características (Pani, 2011): a) o registro de páginas que já foram acessadas por usuários possibilita traçar o seu comportamento, além de prever páginas a serem acessadas; b) o comportamento frequente de acessos a determinadas páginas pode ser usado para identificar a necessidade de criação de *links* para melhorar o desempenho dos acessos futuros; c) baseado no comportamento da navegação de usuários em um *website*, torna-se viável sugerir mudanças no *design* das páginas e; d) os padrões de uso podem ser usados em *Business Intelligence* (BI) com o objetivo de melhorar as vendas e recomendar produtos por meio de propagandas. A Figura 1 apresenta as principais áreas de aplicação da *Web Usage Mining* (Zhang *et al.*, 2009).

## WEB USAGE MINING

Segundo Vellingiri e Pandian (2011), a *Web Usage Mining*, também conhecida como *Web Log Mining*, é um processo de mineração de dados usado para descobrir padrões de uso das informações na Web e tem como objetivo proporcionar o entendimento dos interesses e do comportamento dos usuários web. Assim, analisando os *logs* de acesso de usuários a visitas em *websites*, é possível perceber as suas ações e, conseqüentemente, viabilizar a customização e a personalização para o seu ambiente de navegação.

Geralmente, a *Web Usage Mining* inclui as seguintes etapas (Etminani *et al.*, 2009):

- a. Pré-processamento – remove inconsistências e ruídos das fontes de dados com o objetivo de deixar apenas as informações que são, realmente, significativas. A etapa também inclui tarefas como a identificação do usuário e da sua sessão e a definição do caminho completo da navegação (Rajni e Pramila, 2010).
- b. Descoberta de padrões – consiste em descobrir os interesses de um usuário (ou de uma comunidade) e construir um modelo de acordo com essas preferências.
- c. Análise de padrões – tem como principal objetivo filtrar as informações que, aparentemente são irrelevantes, e, a partir dessa filtragem, visualizar e interpretar os padrões de interesse do usuário.

Enquanto os usuários web navegam nos *websites*, as suas interações são registradas em arquivos chamados de *web log file*. Esse registro tem a forma de uma simples transação e é adicionada a um arquivo de texto ASCII. Atualmente, existem três formas de se obter os *logs* de acesso de um usuário: a) arquivo de *log* de cliente, b) arquivo de *log* de *proxy* e c) arquivo de *log* de servidor. Os delimitadores de um arquivo de *log* podem ser vírgula, espaço em branco ou tabulação.

Existem três tipos de arquivo de *log* de servidor disponíveis para capturar as atividades de um usuário nos *websites* (Hussain *et. al.*, 2010): *Common Log Format*, *Extended Log Format* e *IIS Log Format*. O *Common Log Format* (CLF) é processado por uma variedade de aplicações de servidores web e possui os seguintes campos: *remote host*, *identification*, *authuser*, *date/time*, *HTTP request*, *status code* e *transfer volume*. O *Extended Common Log Format* (ECLF) é uma variação do primeiro e apresenta dois campos a mais: *referrer* e *user agent*. Por fim, o *IIS Log Format* define os campos: *client IP*, *user name*, *date*, *time*, *service and instance*, *server name*, *server IP*, *elapsed time*, *client bytes sent*, *server bytes sent*, *service status code*, *windows status code*, *request type*, *target of operations* e *parameters* (Zdravko e Larose, 2007). Abaixo, são mostrados alguns trechos desses três tipos de arquivos:

```
141.243.1.172 [29:23:53:25] "GET /Software.html HTTP/1.0" 200 1497
query2.lycos.cs.cmu.edu [29:23:53:36] "GET /Consumer.html HTTP/1.0" 200
1325
tanuki.twics.com [29:23:53:53] "GET /News.html HTTP/1.0" 200 1014
wpbfl2-45.gate.net [29:23:54:15] "GET / HTTP/1.0" 200 4889
wpbfl2-45.gate.net [29:23:54:16] "GET /icons/circle_logo_small.gif
HTTP/1.0" 200 2624
wpbfl2-45.gate.net [29:23:54:18] "GET /logos/small_gopher.gif HTTP/1.0"
```

**Figura 2 – Common Log Format**

```
#Software: Microsoft Internet Information Services 6.0
#Version: 1.0
#Date: 2002-05-02 17:42:15
#Fields: date time c-ip cs-username s-ip s-port cs-method cs-uri-stem cs-
uri-query sc-status cs(User-Agent)
2002-05-02 17:42:15 172.22.255.255 - 172.30.255.255 80 GET
/images/picture.jpg - 200
Mozilla/4.0+(compatible;MSIE+5.5;+Windows+2000+Server)
```

**Figura 3 – Extended Log Format**

```
192.168.114.201, -, 03/20/01, 7:55:20, W3SVC2, SERVER, 172.21.13.45, 4502,
163, 3223, 200, 0, GET, /DeptLogo.gif, -,
```

**Figura 4 – IIS Log Format**

Para Eirinaki e Vazirgiannis (2003), a maneira mais comum de analisar um arquivo de *log* de acessos é utilizando um método estatístico. Contudo, existem outros mais sofisticados como Mineração de Regras de Associação, Descoberta de Padrões Sequenciais, *Clustering* e Classificação.

A Mineração de Regras de Associação enfatiza a extração de correlações, padrões frequentes e associações entre registros de bancos de dados ou de outros repositórios. Amplamente utilizada em várias áreas como redes de telecomunicações, gestão de risco e outras (Kotsiantis e Kanellopoulos, 2006). Poderia ser usada também, por exemplo, para mostrar as correlações existentes entre páginas que foram acessadas em uma mesma sessão.

A Descoberta de Padrões Sequenciais é uma extensão da Mineração de Regras de Associação na medida em que são revelados padrões de coocorrência incorporando a noção de sequência de tempo. No domínio da Web, um padrão poderia ser uma página (ou um conjunto delas) que foi acessada após outra página (ou um conjunto). Com essa abordagem, as tendências e as visitas dos usuários podem ser previstas (Eirinaki e Vazirgiannis, 2003).

*Clustering* é um método de mineração orientado por um processo de aprendizado não-supervisionado e consiste no agrupamento de objetos físicos ou abstratos em uma classe formada por objetos similares (Han e Kamber, 2006). Ainda de acordo com as ideias de Eirinaki e Vazirgiannis (2003), no contexto *Web Mining*, existem dois tipos: a) *cluster* de página formado por páginas semelhantes, sob o ponto de vista dos usuários e b) *cluster* de usuário formado por usuários que apresentam comportamentos navegacionais semelhantes.

Por fim, a Classificação é o processo que mapeia um conjunto de dados em uma classe pré-determinada. Ao contrário do *Clustering*, a Classificação é um processo de aprendizado supervisionado e, para isso, necessita que o usuário conheça previamente a definição das classes. Na Web, as classes podem representar diferentes perfis de usuários e utilizam recursos selecionados para descrever a categoria de cada um deles.

## PROPOSTAS DE SISTEMAS DE RECOMENDAÇÃO

Existem várias formas de identificar o interesse de usuários na Web (Xiao *et. al.*, 2001). Neste artigo, serão apresentadas algumas abordagens disponíveis na literatura que se propõem a identificar o comportamento navegacional dos usuários web e a recomendar recursos baseado no interesse de usuários.

O SUGGEST é um sistema baseado em *Web Usage Mining*, proposto por Baraglia e Palmerini (2002), que foi projetado com a finalidade de gerar *hiperlinks* de páginas potencialmente interessantes a usuários. As recomendações são construídas a partir de requisições solicitadas a um servidor web, por meio da análise do comportamento navegacional dos usuários.

Puntheeranurak e Tsuji(2005) utilizaram técnicas de *Web Usage Mining* para construir um protótipo de um sistema de recomendação baseado em palavras-chave e no comportamento navegacional dos usuários. As recomendações são realizadas por meio de perfis de usuários. Uma parte do perfil contém dados sobre os usuários e a outra descreve as suas regras de comportamento.

Tan *et. al.* (2006) propuseram uma abordagem na qual os padrões de navegação descobertos são armazenados em uma estrutura de árvore altamente compactada e são

combinados com um algoritmo de recomendação sem, necessariamente, gerar regras de associação.

O trabalho de Esslimani *et. al.* (2008) apresenta um sistema de recomendação híbrido que combina perfis de usuários à abordagem colaborativa. Os autores propõem uma nova métrica para o cálculo da similaridade navegacional considerando as sequências de navegação do usuário ativo com as dos demais. Dessa forma, o sistema faz recomendações usando essas correlações e métodos de filtragem colaborativa. O objetivo da combinação é avaliar o impacto de cada técnica na acurácia das recomendações.

A maioria dos trabalhos na área de personalização web faz apenas a recomendação de páginas com conteúdos interessantes para o usuário que está navegando, porém, na prática, a sua necessidade de interesse é mais que disso. Lei e Fan (2008) abordaram um método de personalização que, além de recomendar páginas interessantes aos usuários, pode descobrir seus potenciais domínios de interesse. A abordagem também possibilita que os usuários encontrem páginas relevantes dentro do domínio selecionado.

*NetPersonal* é um sistema de recomendação proposto por Thakur *et. al.* (2009) com o objetivo de personalizar o ambiente de usuários web. A ferramenta é inferida pelos *logs* de acesso nos servidores web, obtidos por meio do uso de técnicas de *Web Mining* para extrair conhecimento dos perfis de usuários. O algoritmo de recomendação é baseado em *clusters* que, essencialmente, representam um grupo de usuários com padrões de acessos similares.

De modo geral, as técnicas de WUM produzem padrões representados por páginas web e, dessa forma, o significado semântico do conteúdo visitado não é incluído nos processos de mineração de uso web. Assim, Salin e Senkul (2009) apresentaram um *framework* para integrar informações semânticas às técnicas de *Web Usage Mining* a fim de enriquecer os padrões de navegação minerados. Os padrões são extraídos na forma de instâncias de ontologias, ao invés de páginas web, e os resultados obtidos por elas são usados para fazer a recomendação de páginas.

Wei e Shu-hai (2009) apresentaram um modelo híbrido, que integra agrupamento de páginas web às técnicas de *Web Usage Mining*, para descoberta e análise dos padrões de navegação. Os resultados obtidos indicam que a integração de páginas agrupadas, conhecimento semântico e mineração de uso podem produzir resultados satisfatórios aos sistemas de recomendação.

Wang *et. al.* (2010) propuseram um sistema de recomendação que objetiva identificar as páginas web que apresentam a maior probabilidade de serem acessadas por um determinado usuário. Para isso, a abordagem utiliza técnicas de mineração de padrões sequenciais, bastante úteis para predição, e ainda propõe um modelo compacto de dados chamado *Pattern-tree*, o qual armazena a sequência dos padrões de acesso.

Os interesses de usuários web são imprecisos e incertos, por isso as técnicas que usam a lógica *booleana* para fazer recomendações tornam-se inadequadas. Nesse contexto, os algoritmos de agrupamento *fuzzy* aparecem como uma boa alternativa para tal propósito. O trabalho de Nadi *et. al.* (2010), baseado em técnicas de agrupamento *fuzzy*, é um modelo de recomendação que combina a *Web Content Mining* com a *Web Usage Mining* a fim de encontrar regras de interesse. As técnicas de agrupamento *fuzzy* foram usadas para agrupar cada usuário de

acordo com o conhecimento prévio de suas preferências e também para possibilitar a identificação das incertezas no seu comportamento.

Takano e Li (2010) apresentaram um sistema de recomendação de páginas para *e-learning*, o qual utiliza um método híbrido de *feedback* para extrair as preferências e o comportamento navegacional dos usuários. O modelo proposto adapta-se às preferências individuais do usuário e também ao seu interesse de querer mudar a sua atividade de aprendizagem. Os autores apontam que uma das vantagens do *feedback* ser implícito é o fato de ele ser efetuado de forma transparente para o usuário.

Muitas pesquisas relacionadas à aprendizagem personalizada vêm sendo realizadas e, por consequência, algumas ferramentas foram propostas como, por exemplo, ATutor, WebCT, BlackBoard e outras. Qi *et. al.* (2010) apresentaram o Sistema de Recomendação de Recurso Personalizado para *e-Learning*, que utiliza técnicas de *Data Mining*, para recomendar recursos personalizados a alunos que utilizam plataformas *e-Learning*. O sistema possui quatro módulos (modelo do estudante, agrupamento de materiais para aprendizagem, recomendações de materiais e avaliação dos materiais recomendados) e baseia-se nas características dos usuários e em suas preferências de aprendizagem e, assim, é feita a recomendação de materiais personalizados para aprendizagem.

No trabalho de Shan e Ren (2010), é proposto o *Browsing Behavior Personalized Information Recommendation System* (BBIRS), também baseado no comportamento de navegação do usuário, que objetiva recomendar recursos de *e-Learning*. Contudo, a recomendação é feita sob a perspectiva de que o grau de interesse de um usuário ( $Interest(p) = \frac{1}{\max(\frac{1}{Speed(p)}, \frac{1}{Speed(v)})}$ ), em relação a uma determinada página, é calculado de acordo com a velocidade de navegação ( $Speed(p) = \frac{Size(p)}{Duration(p)}$ ). A acurácia do BBIRS e a do modelo de Markov foram comparadas e constatou-se que a performance da proposta dos autores foi superior, indicando que o sistema pode ajustar-se às mudanças de interesse dos usuários.

Os padrões de acessos periódicos ocorrem, frequentemente, em um determinado período do tempo, por exemplo, das 14h às 15h. A proposta de Fong *et. al.* (2011) utiliza a representação da lógica *fuzzy* para construir uma base de conhecimento que é atualizada por meio dos padrões de acessos periódicos. A partir das informações armazenadas na base de conhecimento, os hábitos e o comportamento dos usuários são modelados para que sejam feitas recomendações.

## CONCLUSÕES

A popularização da Internet causou um impacto no número de páginas publicadas na rede. Por um lado, a quantidade de informações disponíveis configura uma imensa fonte de dados, por outro, a dificuldade encontrada pelos usuários web para localizar recursos que atendam aos seus interesses é um problema que algumas áreas, como a Recuperação da Informação e a Mineração na Web, vêm tentando resolver.

Nos últimos anos, diversos trabalhos apresentaram propostas para sistemas de recomendação baseados no perfil de interesse do usuário, apesar de essa identificação não ser

uma tarefa fácil. Nesta pesquisa, foram apresentadas 14 (catorze) propostas de sistemas de recomendação que utilizam técnicas da *Web Usage Mining* e o objetivo foi elencar algumas características de cada uma delas, além de estimular o desenvolvimento de novos trabalhos nesta área.

## REFERÊNCIAS BIBLIOGRÁFICAS

1. BARAGLIA, R., PALMERINI, P. SUGGEST: A Web Usage Mining System. *Information Technology: Coding and Computing*, p. 282- 287, abr. 2002.
2. EIRINAKI, M, VAZIRGIANNIS, M. Web Mining for Web Personalization. *ACM Transactions on Internet Technology (TOIT)*, v. 3, p. 1-27, 2003.
3. ESSLIMANI, I., BRUN, A., BOYER, A. Enhancing collaborative filtering by frequent usage patterns. *Applications of Digital Information and Web Technologies, 2008. ICADIWT 2008. First International Conference*, p. 180-185, ago. 2008.
4. ETMINANI, K.; DELUI, A.R.; YANEHSARI, N.R.; ROUHANI, M. Web usage mining: Discovery of the users' navigational patterns using SOM. *Networked Digital Technologies, 2009. NDT '09. First International Conference*, p. 224-249, jul. 2009.
5. FONG, A.C.M.; ZHOU, B., HUI, S.C.; HONG, G.Y., DO, T.A. Web content recommender system based on consumer behavior modeling. *Consumer Electronics, IEEE Transactions*, v.57, n.2, p. 962-969, mai. 2011.
6. HAN, J., KAMBER, M. *Data Mining: Concepts and Techniques*. Morgan Kaufmann, 2nd edition, 2006.
7. HUSSAIN, T., ASGHAR, S., MASOOD, N. Web usage mining: A survey on preprocessing of web log file. *Information and Emerging Technologies (ICIET), 2010 International Conference*, p. 1-6, jun. 2010.
8. KOTSIANTIS, S., KANELLOPOULOS, D. Association Rules Mining: A Recent Overview. *GESTS International Transactions on Computer Science and Engineering*, 2006, v.32, n. 1, p. 71-82.
9. LEI, M., FAN, L. A Web Personalization System Based on Users' Interested Domains. *Cognitive Informatics, 2008. ICCI 2008. 7th IEEE International Conference*, p.153-159, ago. 2008
10. LING, H., LIU, Y., YANG S. An Ant Colony Model for Dynamic Mining of Users Interest Navigation Patterns. *Control and Automation. 2007 IEEE International Conference*, p. 281-283, 2007.
11. MARKOV, Z., LAROSE, D. T. *Data Mining the Web – uncovering patterns in Web Context, Structure, and Usage*, Wiley, New Jersey, 2007.
12. NADI, S., SARAEE, M., DAVARPANAH-JAZI, M. A fuzzy recommender system for dynamic prediction of user's behavior. *Internet Technology and Secured Transactions (ICITST), 2010 International Conference*, p.1-5, nov. 2010.
13. PANI, S. K. *et al.* Web Usage Mining: A Survey on Pattern Extraction from Web Logs. *International Journal of Instrumentation, Control & Automation*, v.1, n.1, 2011.

14. PAMNANI, R., CHAWAN, P., Web Usage Mining: A Research Area in Web Mining. In: Proceedings of IS CET 2010, p. 73-77, 2010.
15. PUNTHEERANURAK, S., TSUJI, H. Mining Web logs for a personalized recommender system. Information Technology: Research and Education, 2005. ITRE 2005. 3rd International Conference, p. 445- 448, jun. 2005.
16. QI, H., CUI, M., XIAO, M. A Personalized Resource Recommendation System Using Data Mining. E-Business and E-Government (ICEE), 2010 International Conference, p. 5365-5368, mai. 2010.
17. RAJNI, P., PRAMILA, C. Web Usage Mining: A Research Area in Web Mining. In Proceedings of IS CET 2010. p. 73-77, 2010.
18. SALIN, S., SENKUL, P. Using semantic information for web usage mining based recommendation. Computer and Information Sciences, 2009. ISCIS 2009. 24th International Symposium, p. 236-241, set. 2009.
19. SHAN, R., REN, Z. Research on personalized recommendation system in E-learning. Education Technology and Computer (ICETC), 2010 2nd International Conference, vol.4, p. v.4-182- 184, jun. 2010.
20. STUMME, G., HOTHO, A., BERENDT, B. Semantic Web Mining – State of the art and future directions. Journal of Web Semantics, p. 124-143, 2006.
21. TAKANO, K., LI, K.F. An Adaptive e-Learning Recommender Based on User's Web-Browsing Behavior. P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC), 2010 International Conference, p. 123-131, nov. 2010.
22. TAN, X., YAO, M., XU, M. An Effective Technique for Personalization Recommendation Based on Access Sequential Patterns. Services Computing, 2006. APSCC'06. IEEE Asia-Pacific Conference, p.42-46, dez. 2006.
23. THAKUR, B.K., ABBAS, S.Q., TRIVEDI, A.K. NetPersonal: A Recommender System to Personalize the Environment of Web User. Advance Computing Conference, 2009. IACC 2009. IEEE International, p.702-705, mar. 2009.
24. VELLINGIRI, J., PANDIAN, S. C. A Survey on Web Usage Mining. Global Journal of Computer Science and Technology, p. 66-72, 2011.
25. WANG, X., BAI, Y., LI, Y. An Information Retrieval Method Based On Sequential Access Patterns. Wearable Computing Systems (APWCS), 2010 Asia-Pacific Conference, p. 247-250, abr. 2010.
26. WEI, L., SHU-WAI, Z. A Hybrid Recommender System Combining Web Page Cluster. Computational Intelligence and Software Engineering, 2009. CiSE 2009. International Conference, p. 1-4, dez. 2009.
27. XIAO, J., ZHANG, Y., JIA, X., LI, T. Measuring similarity of interests for clustering Web-users. Database Conference, ADC 2001, p. 107-114, 2001.
28. ZHANG, J., ZHAO, P., SHANG, L., WANG, L. Web usage mining based on fuzzy clustering in identifying target group. Computing, Communication, Control and Management, 2009. CCCM 2009. ISECS International Colloquium, v. 4, p.209-212, ago. 2009.