



REPOSITÓRIO DE VÍDEOS COM SUPORTE A RECOMENDAÇÃO DE INFORMAÇÕES

ZANCHI, Fábio da Silva¹; CHICON, Patricia Mariotto Mozzaquatro²

Palavras-Chave: Repositório. Extração da informação. Recomendação da informação.

INTRODUÇÃO

O crescente avanço e evolução da tecnologia da informação possibilitou que grandes volumes de dados possam ser armazenados e processados mais rapidamente e com maior precisão (SANTOS, 2002).

Com o aumento da disponibilidade de informações, torna-se cada vez mais difícil para um usuário encontrar o que realmente é de seu interesse. Uma procura por um produto, uma página, um serviço ou qualquer outro tipo de informação geralmente é muito difícil, por retornar uma lista de itens encontrados extensa e com vários itens irrelevantes.

Como solução a este problema pode-se criar um sistema que pudesse reconhecer o interesse do usuário e o auxiliar em sua consulta. Esse sistema filtraria os itens irrelevantes e classificaria os itens encontrados por ordem de relevância, além de poder sugerir itens que provavelmente são do interesse do usuário, mas não satisfazem sua consulta. Esse tipo de sistema é denominado Sistema de Recomendação.

O resumo apresentado é parte de trabalho de conclusão de curso que objetiva implementar um repositório de vídeo com suporte a recomendação de informação.

EXTRAÇÃO DA INFORMAÇÃO E RECUPERAÇÃO DA INFORMAÇÃO

De acordo com Gaizauskas e Wilks (1998), Extração de Informação é o termo que se dá a atividade de extrair automaticamente uma informação de um tipo de texto de língua natural. A informação extraída é determinada por um conjunto de padrões ou regras de extração

¹ Discente do Curso de Ciência da Computação. UNICRUZ. E-mail: fzanchi@unicruz.edu.br

² Professora, UNICRUZ. E-mail: pmozzaquatro@unicruz.edu.br



específicas ao seu domínio: padrões que podem ser escolhidos de maneira manual, por algum especialista, ou de forma automatizada.

A área de Extração de Informação visa localizar e extrair, de forma automática, informações relevantes em um documento ou coleção de documentos expressos em língua natural e estruturar tais informações para os padrões de saída, por exemplo, em banco de dados, a fim de facilitar sua manipulação e análise (Grishman, 1997).

Sistemas em Extração de Informação extraem tipos específicos de informações a partir de bases de dados não estruturadas. Sua vantagem é a segmentação do texto de entrada, permitindo que partes não pertinentes ao domínio, como frases ou orações inteiras, sejam ignoradas. Isso simplifica consideravelmente o processamento, torna-o menos oneroso e reduz a ocorrência de problemas difíceis, como a resolução de ambiguidades.

Recuperação de Informação (RI) é a área de pesquisa que trata da representação, armazenamento e acesso a itens de informação. O problema principal em um sistema de recuperação é encontrar o subconjunto de objetos relevantes à consulta especificada pelo usuário. A representação e o armazenamento devem prover acesso fácil a essas informações (BAEZA-YATES & RIBEIRO-NETO, 1999).

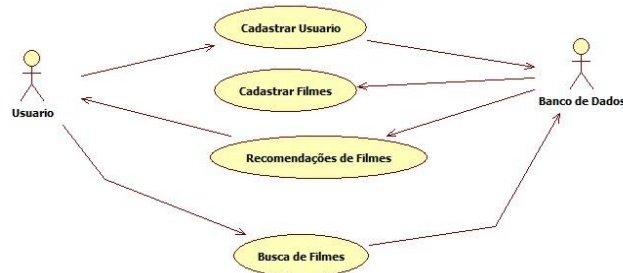
Belvin (1992 apud CAZELLA, 2006) descreve um Sistema de Recuperação como um sistema que conduz o usuário diretamente para aquilo que lhe é interessante. O autor cita que os exemplos mais encontrados são: os motores de busca na *Web* e os *sites* de pesquisa.

METODOLOGIA E DESENVOLVIMENTO

Na Figura 1 pode ser observada o Diagrama de Caso de Uso que representa o repositório de vídeo. Primeiramente, pode-se observar dois atores, o usuário e o banco de dados, sendo que o primeiro ao fazer o cadastro na qual escolhe o tipo de gênero que gosta, pode interagir na busca por um filme e o segundo é onde se faz o cadastramento dos filmes e as recomendações de filmes para os usuários.



Figura 1 – Diagrama de Caso de Uso

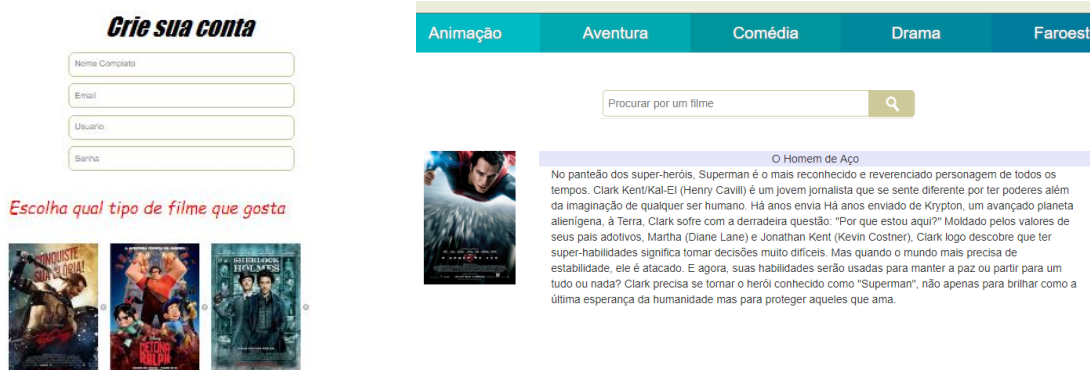


Fonte: Elaborado pelo Autor

O desenvolvimento foi realizado nas seguintes etapas: Criação do layout e linguagem de programação, Procura pelas informações, Inserção no banco de dados, Desenvolvimento do sistema³. A Figura 2 mostra a tela de cadastro, como também o processo de pesquisa de um filme realizado.

Figura 2

- Cadastro e pesquisa



Fonte: Elaborado pelo Autor

Na Figura 2 tem-se a tela de cadastro, na qual o usuário irá preencher os dados e escolher qual tipo de filme preferencial. Também ilustra o resultado de uma pesquisa a qual sugere a partir do seu perfil cadastrado, filmes de ação.

CONSIDERAÇÕES

Nos últimos anos são evidentes os avanços da Tecnologia de Informação (TI) e o intenso aumento do volume de dados que chegam as pessoas a todo momento.

³ Disponível em: <http://wysemovie.epizy.com/>



A dificuldade de encontrar informações relevantes e confiáveis em buscas na web e com o desenvolvimento de um repositório de vídeo, o qual tornará mais ágil a procura com maior qualidade e precisão na pesquisa solicitada.

Este resumo é parte integrante de um trabalho de conclusão de curso em andamento que objetiva a implementação das técnicas de filtragem baseada em conteúdo e filtragem colaborativa no suporte a recomendação de vídeos analisando a precisão na extração e recuperação de informações conforme o perfil do usuário.

REFERÊNCIAS

BAEZA-YATES, R. A. e B. A. Ribeiro-Neto. **Modern Information Retrieval: ACM Press, Addison-Wesley.** 1999. 513 p.

BELVIN, N. J. CROFT, W. B. **Information Filtering and Information Retrieval: two sides of the same coin?** Communications of the ACM, New York, v.35, n.12, p. 29, Dez. 1992.

GAIZAUSKAS, R.; WILKS, Y. **Information extraction: Beyond document retrieval.** Journal of Documentation, v. 54, n. 1, p. 70-105, 1998.

GRISHMAN, R. **Information extraction: Techniques and challenges.**In SCIE '97: International Summer School on Information Extraction, pp. 10–27. Springer-Verlag. 1997.

SANTOS, Maria de Fátima Almeida. Integração - **Mineração de Dados – SGBD não é uma panacéia: Estudo da Integração do Algoritmo Apriori Quantitativo ao Oracle9I.** Dissertação de Mestrado, Universidade Federal de Campina Grande, Campina Grande, 2002.