

Personal_Movie – Um modelo de Sistema de Recomendação de filmes geolocalizados em eventos.

Otávio Cordeiro Siqueira de Oliveira (Universidade Federal de Sergipe, Sergipe, Brasil) – otasiq@gmail.com

Maria Augusta Silveira Netto Nunes (Universidade Federal de Sergipe, Sergipe, Brasil) - gutanunes@gmail.com

Silvio César Cazella. (Universidade do Vale do Rio dos Sinos, Rio Grande do Sul, Brasil) - cazella.unisinos@gmail.com

Resumo — Este trabalho descreve um modelo que aplica informações contextuais em Sistemas de Recomendação para predição de filmes geolocalizados. Este modelo visa minimizar o problema de sobrecarga de informação e a dificuldade da disponibilização da informação personalizada que é intensificado em grandes eventos como a Copa do Mundo, os Jogos Olímpicos e os Jogos Paraolímpicos. Para a análise do modelo proposto foi desenvolvido um protótipo que utiliza a geolocalização para obtenção de informações contextuais que será disponibilizado para testes com usuários finais a fim validar o seu potencial de assertividade na geração de recomendação.

Palavras-chave - Sistemas de Recomendação, Geolocalização, Predição, Informações Contextuais, Eventos.

Personal_Movie - A Model System Recommendation geolocalizados film events.

Abstract — *This paper describes a model that applies contextual information on Recommender Systems for predicting geolocalizados films. This model aims to minimize the problem of information overload and the difficulty of providing personalized information that is heightened in big events like the World Cup, the Olympic and Paralympic Games. For the analysis of the proposed model was developed a prototype that uses geolocation to obtain contextual information which will be available for testing with end users to validate its potential in the generation of assertiveness recommendation.*

Keywords - Recommender Systems, Geolocation, Prediction, contextual information, events.

Agradecimentos: Agradeço ao Programa Institucional de Bolsas de Iniciação Científica (PIBIC) que apoiou a pesquisa.

1 Introdução

Nos próximos anos, o Brasil sediará a Copa do Mundo, os Jogos Olímpicos e os Jogos Paraolímpicos e com o transcorrer desses eventos surge uma demanda maior por informações personalizadas sobre as atrações ofertadas pelos eventos como filmes, porém muitas vezes a informação não é disponibilizada de forma personalizada aos telespectadores, fazendo com que muitos deles possam receber uma sobrecarga de informação e dificultar a escolha de uma atração.

Essa sobrecarga de informações, ou seja, à quantidade de informação que estava sendo gerada pelos diversos tipos de sistemas e recebida pelos usuários já era uma preocupação dos autores Loeb e Terry [1]. Com o propósito de suprir essa preocupação surgiram os Sistemas de Recomendação (SR) que auxiliam no aumento da capacidade e eficácia do processo de indicação de itens para usuários [2].

Esses sistemas fornecem recomendações sobre itens para usuários com potencial de aceitação, ou seja, realiza a união adequada entre o que está sendo recomendado e os que estão recebendo a recomendação. Esse processo de recomendar é um dos grandes desafios desse tipo de sistema.

A inserção de informações contextuais como: horário do filme, local de exibição, preço do bilhete do cinema e a censura auxiliam o processo de recomendação refinando-a. Com essas informações é possível prever itens de acordo com o contexto no qual o usuário está inserido [3].

O uso de recursos presentes em dispositivos móveis como a Geolocalização atrelado aos recursos dos SR pode vir a auxiliar a obtenção de informações contextuais de maneira eficiente, e pode auxiliar a disponibilização da recomendação de forma eficaz.

Nas sessões a seguir serão apresentadas características dos SR's, técnicas de recomendação, trabalhos relacionados, detalhes do protótipo e a conclusão parcial em conjunto com os trabalhos futuros.

2 SISTEMAS DE RECOMENDAÇÃO

Em Sistemas de Recomendação, normalmente as pessoas baseiam-se nas recomendações de amigos, através de opiniões de especialistas, ou ainda, a partir de outras fontes para tomar uma decisão. Os Sistemas de Recomendação ajudam e aumentam este processo social natural já existente entre as pessoas [23].

Para Resnick e Varian o grande desafio está em descobrir o relacionamento de interesses entre os usuários, buscando prever os itens de maior interesse para ele [2].

Para prever itens é necessário obter de dados sobre os mesmos. A coleta de dados pode ser realizada de forma explícita ou implícita. Gadhane e Lhuillier [4] afirmam que a coleta explícita sobre produtos é aquela em que o usuário necessita inferir manualmente suas preferências sobre itens, porém nessa modalidade de

coleta o usuário pode não especificar seus interesses completamente, podendo assim mascarar possíveis resultados.

A coleta de dados implícita utiliza técnicas extração de conteúdo, mineração de dados, entre outras, para que se possam coletar dados do cliente sem que o mesmo perceba, nesse modelo se extrai informações de possíveis áreas de interesse e aplicando técnicas associação é possível sugerir produtos e serviços [4].

A coleta implícita também deverá auxiliar a obtenção dos dados do usuário de modo que possa vir a formar seu perfil [18]. Segundo os autores Cazella, Nunes e Reategui [16] quanto maior o nível de informações sobre o usuário, mais pertinentes serão as recomendações de produtos, serviços e/ou pessoas. Para formação do perfil do usuário se faz necessário definir a personalidade do indivíduo e para isso, existe uma abordagem bastante interessante sobre o assunto que é a de traços de personalidade que permite diferenciar psicologicamente pessoas usando traços mensuráveis e conceituáveis [17].

Após a coleta de dados aplicam-se técnicas para recomendar esses itens coletados, na sessão a seguir serão apresentadas as principais técnicas para a geração da recomendação.

2.1 Técnicas de Recomendação

Técnicas de filtragem de informação possuem uma interação dinâmica com os usuários, a cada nova ação realizada o algoritmo deve procurar um novo item que atenda ao usuário naquele momento e não precisa ser iniciado pelo mesmo, a seguir serão citadas as principais técnicas para recomendação.

Na filtragem baseada em conteúdo (FBC) baseia-se exclusivamente nas variáveis usuário e item, comparando o conteúdo dos itens às preferências dos usuários. A categorização dos itens é o grande problema desta técnica, a busca de informação enfrenta um problema muito comum em textos que é o sentido de um metadado.

O tratamento de metadados é um grande desafio para a técnica. O processo de identificação do conteúdo dos atributos e diferenciação do sentido dele naquele contexto específico e um grande desafio em determinado contexto [5].

A filtragem colaborativa (FC) utiliza as avaliações realizadas pelos usuários aos itens e as compara com as avaliações realizadas por outros usuários.

Os autores Adomacius e Tuzhilin [5] destacam que uma característica importante da FC e que nesse modelo utiliza uma técnica de recomendação na qual independe de domínio. Ela é indicada para recomendação de conteúdo que não pode ser escrito adequadamente por metadados.

O funcionamento da FC divide-se basicamente em três etapas [6] [7]:

1. Cálculo do peso de cada usuário em relação à similaridade ao usuário ativo
2. Selecionar um subconjunto de usuários com maiores similaridades (vizinhos) para considerar na predição.
3. Normalizar as avaliações e computar as predições ponderando as avaliações dos vizinhos com seus pesos.

Essa técnica também é chamada de “*k-nearest-neighbor*” ou “*user-based*” [5].

A Filtragem Baseada em Informação Contextual (FBIC) [9] surgiu para suprir a lacuna deixada pelas abordagens tradicionais 2D citadas anteriormente, onde baseiam-se apenas em Usuário X Item e não tratando informações contextuais.

O conceito de contexto está inserido de forma implícita nas atividades corriqueiras dos seres humanos e ajuda na comunicação. Quando o contexto é percebido permite a uma pessoa fazer avaliações, tomar decisões e adaptar seu comportamento de acordo com a situação. Para que uma pessoa possa tomar decisões de forma apropriada à situação se faz necessário compreender o contexto [8].

Já a Filtragem Híbrida (FH) faz uso de umas ou mais técnicas de recomendação, geralmente combinando uma abordagem colaborativa baseada em conteúdo com a filtragem colaborativa. Pode-se ainda utilizar mais que técnica para a filtragem baseada em conteúdo.

A FH vem sendo amplamente utilizada em Sistemas de Recomendação, pois utilizam mais de uma técnica de recomendação conjugada para prevenir problemas clássicos de “*cold start*”, que é o problema da falta de informações necessária para gerar a recomendação.

4 Trabalhos Relacionados

Vários trabalhos vêm abordando diferentes propostas para a recomendação de produtos como filmes. O NetFlix [24] propõe a utilização de mineração de dados e filtragem baseada em conteúdo, além da coleta de informação de forma implícita, na qual o sistema infere as preferências dos usuários.

Já o MovieLens [20] e o IMDB [18] fazem uso da FBC e a coleta explícita do usuário na qual o mesmo deve inferir suas preferências sobre itens, além de usar técnicas de FC para melhorar o seu processo de recomendação [10].

Pesquisadores a fim de aumentar a eficácia dos SR têm utilizado a coleta implícita em conjunto com a explícita onde o usuário informa manualmente suas preferências através de questionários, aplicando notas a itens ou selecionando áreas de interesse pessoal. É o caso do Jinni [25], HunchMovies [26], WhattoRent [21].

A técnica utilizando informações contextuais está presente nos SR’s Alfred e DITTO [22] em conjunto com a técnica de filtragem baseada em conteúdo. Esses

SR’s não recomendam filmes, apenas serviços para usuário utilizando dados do GPS para realizar a coleta de informações contextuais.

O CinemaKI [11] apesar de coletar informações contextuais do usuário nele é possível recomendar filmes que estão em cartaz nos cinemas de acordo com a coleta explícita realizada pelo serviço, além de usar a FC E FBC para predizerem itens aos usuários.

A seguir, no quadro 1, é apresentado um comparativo entre os trabalhos relacionados aqui expostos e o modelo do Personal_Movie proposto neste trabalho. Foram elencadas algumas características e marcado com X a presença dela no sistema estudado. Foi avaliado o método de coleta de informação, a utilização de mineração de dados, a técnica de filtragem de informação, e se o sistema utiliza ou não informações contextuais e Geolocalização.

	IMDB	MovieLens	NETFLIX	CinemaKI	Jinni	HunchMovies	WhattoRent	Alfred	DITTO	Personal_Movie
Mineração de Dados			x							
Coleta Implícita			x	x		x		x	x	X
Coleta Explícita	x	x	x	x	x	x	x	x	x	X
FC	x	x	x	x	x	x	x	x	x	X
FBC	x	x	x	x	x	x	x			X
Informações Contextuais								x	x	X
Geolocalização								x	x	X

Quadro 1 - Comparativo entre os trabalhos relacionados.

Nas sessões a seguir serão apresentados detalhes da modelagem do protótipo, assim como as suas principais funcionalidades.

5 Personal_Movie

O modelo proposto por este trabalho servirá de base para a construção de SR de filmes geolocalizados, unindo informações contextuais a técnicas de recomendação, para modelo o filme é tratado como um produto podendo assim ser adaptado para qualquer produto na recomendação.

As informações contextuais ajudam a recuperar filmes que mais se adaptem ao contexto no qual o usuário está inserido. Utilizando informações contextuais na fase de pré-filtragem coletam-se informações de modo implícito sobre filmes que estão em cartaz nos cinemas é que se situam dentro contexto atual do usuário, posterior a fase de coleta realiza-se a filtragem baseada em conteúdo para categorização dos filmes e armazenamento dos dados sobre programação do cinema.

Utilizando um perfil mais rico do usuário com a

agregação sobre as interações realizadas no serviço ao cálculo de similaridade entre usuários incluem-se estas informações no processo de filtragem colaborativa. Desta forma, este processo seleciona a vizinhança que possui as avaliações mais similares ao usuário alvo.

Após finalizar a nota predita é realizado o ajuste na fase de pós-filtragem da seguinte maneira: caso o contexto do usuário possui o gênero do filme a ser avaliada, a nota predita é acrescida de 30% e em caso contrário, decrescida de 30%. Da mesma forma, caso o contexto possua esta classificação cadastrada, a nota é acrescida de 20% e se não, decrescida de 40%. As avaliações realizadas pelos usuários de itens da base de dados são definidas utilizando a escala Likert [27] de cinco pontos. As recomendações geradas também usam escala Likert de cinco pontos e somente devem ser dirigidas aos usuários caso a nota predita seja igual ou superior a três.

5.1 Protótipo

O Protótipo Personal_Movie utiliza conceito de SOA (Service-Oriented Architecture) que é uma abordagem para criar sistemas computacionais distribuídos baseados no encapsulamento de funções de negócio em serviços que podem ser acessados de forma fracamente acoplada [15].

O Personal_Movie dispõe de um serviço web que contém a lógica de negócio para o processo de recomendação de filmes geolocalizadas. Através da interface do dispositivo móvel é possível coletar informações contextuais e envia um servidor HTTP que contém o *Web Service* para processá-las e disponibilizar a recomendação para o usuário. Ele pode identificar o quando o filme recomendado atende às suas necessidades ou desejos do usuário. Esta informação alimenta o sistema como uma nova avaliação, aumentando a precisão de futuras sugestões.

O contexto cujo usuário está inserido no momento da recomendação é informado de forma implícita através das coordenadas de localização obtidas pelos dados do GPS, porém a interface é configurável para simular obtenção de informações contextuais em localidade distintas.

Após a coleta de informações contextuais o motor de recomendação aplica técnicas de FBC para categorização das informações sobre filmes e realizar o armazenamento dessas informações na base de dados usando o SGBD SQL Server 2008 R2.

Para recomendação o protótipo usa a técnica de recomendação FC, na qual calcula a similaridade entre itens usando a Correlação de Pearson, para que seja possível calcular a similaridade entre itens dentre os quais o usuário-alvo votou.

Cada avaliação realizada pelo usuário pelo Personal_Movie é enviada para o servidor e armazenada no sistema. Quando o algoritmo de recomendação for executado novamente por qualquer usuário do sistema, estas informações transmitidas no *feedback* realizado

serão processadas e poderão gerar uma recomendação diferente das anteriormente criadas.

O Personal_Movie utiliza também conceitos de Design Patterns, pois propõe um padrão para recomendação de filmes em cartaz nos cinemas que independente da interface de apresentação, o serviço de recomendação usa o padrão *Model-view-controller* (MVC), onde as regras de negócio estarão no *Web Service* e a visão e o controle é realizadas através do celulares que possuem a plataforma Android [17].

Através do dispositivo móvel do próprio o usuário do serviço o mesmo poderá receber recomendação de maneira rápida e eficaz, podendo auxiliar o processo natural de escolha de filmes com maior possibilidade de assertividade na escolha, como exemplifica o protótipo de tela exposto na figura 1.



Figura 1 – Recomendação de Filmes

A interface do protótipo permite também a visualização de contatos que estão próximos ao usuário-alvo e estão vinculados a sua conta do Gmail. A funcionalidade permite que um usuário adicione um novo contato e vincule-o aos seus contatos da Google, também é possível registrar a localização atual de um usuário para possa ser visualizada pelos seus contatos. Com essa funcionalidade é possível que um usuário desperte o interesse de convidar amigos para assistir um filme em conjunto com o mesmo.

6 CONCLUSÃO PARCIAL E TRABALHOS FUTUROS

Até o momento foi desenvolvido o serviço web e a interface com o dispositivo móvel para que os turistas dos eventos esportivos possam receber informações personalizadas sobre as atrações *in loco*. Com os resultados obtidos até o presente momento, é possível vislumbrar a aplicação de técnicas para recomendação de grupos e técnicas de personalidade que segundo Nunes [14] pode melhorar a recomendação, para que

um usuário que deseja assistir a uma atração do evento esportivo receba a recomendação de quais amigos ele pode convidar para assistir o mesmo de acordo com os traços de personalidade dos indivíduos. O protótipo também irá sofrer tradução para as línguas inglesa e espanhola, a fim de facilitar a interação com os turistas presentes nos jogos Olímpicos, Paraolímpico e na Copa do Mundo.

REFERÊNCIAS

- [1] Loeb, S.; Terry, D. (1992). Information Filtering. *Communications of ACM*, New York, v.35, n.12, p.26, Dec.
- [2] Resnick, Paul e Varian, Hal R. Recommender Systems. In: *Communications of the ACM*, 1997, Nova Iorque, p. 56-58.
- [3] Ricci, F.; Rokach, L.; Shapira, B.; Kantor, P.B. *Recommender Systems Handbook* 1st Edition., 2011, XXIX, 842 p. 20 illus.
- [4] Gadanho, Sandra C. e Lhuillier, Nicolas. Addressing uncertainty in implicit preferences. In: *RECSYS '07: ACM Conference on Recommendation Systems*, 2007, Minneapolis, p. 97-104.
- [5] Adomavicius, Gediminas e Tuzhilin, Alexander. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. In: *IEEE Transactions on Knowledge and Data Engineering*, 2005, Los Alamitos, p. 734-749.
- [6] Adomavicius, Gediminas e Tuzhilin, Alexander. Context-Aware Recommender Systems. In: *RICCI*, Francesco et al. *Recommender Systems Handbook*. Nova Iorque: Springer, 2011, cap. 7, p. 217-253.
- [7] Adomavicius, Gediminas et al. Incorporating contextual information in recommender systems using a multidimensional approach. In: *ACM Transactions on Information Systems*, 2005, Nova Iorque, p. 103-145.
- [8] Vieira, V.; Souza, D. Salgado, A.C.; Tedesco, P. Uso e Representação de Contexto em Sistemas de Computacionais, Mini-curso apresentado no Simpósio de Fatores Humanos em Sistemas Computacionais (IHC 2006), Natal, Brasil, 2006.
- [9] Vieira, V., Tedesco, P., Salgado, A. C. (2009) "A Process for the Design of Context-Sensitive Systems", In: *Proc. of the 13th International Conference on Computer Supported Cooperative Work in Design (CSCWD'09)*, pp. 143-148, Santiago, Chile.
- [10] Dhirosattha. Predictive algorithm and what's so great about the Netflix Prize, 2009. Disponível em <<http://sendy82.wordpress.com/2009/09/22/predictive-algorithm-and-whats-so-great-about-the-netflix-prize/>>. Acesso em 07/02/2012.
- [11] Idg Now, IG lança rede social focada em cinema, 2011. Disponível em <<http://idgnow.uol.com.br/internet/2010/10/20/ig-lanca-rede-social-focada-em-cinema/>>. Acessado em 14/01/2012.
- [12] Google. Android Developers, 2012. Disponível em <<http://developer.android.com>>. Acessado em 10/12/ 2011.
- [13] Morado, Denise et al. GESTÃO ESTRATÉGICA DA INFORMAÇÃO: a distribuição da informação e do conhecimento, 2008. Disponível em <http://www.arquivar.com.br/espaco_profissional/sala_leitura/artigos/gestao_estrategica_info.pdf/view?searchterm=None>. Acessado em 05/02/2012.
- [14] Nunes, Maria A. S. N. Recommender Systems based on Personality Traits. Tese. Universidade Montpellier 2, Montpellier, 2008.
- [15] ERL, Thomas. Service-Oriented Architecture. Concepts, Technology, and Desing. Indiana, Estados Unidos: Prentice Hall, 2005.
- [16] CAZELLA, Sílvio C.; NUNES, Maria A. S. N.; REATEGUI, Eliseo B. A Ciência do Palpite: Estado da Arte em Sistemas de Recomendação. In: *Jornada de Atualização de Informática – JAI 2010 – CSBC*. Rio de Janeiro: PUC RIO, v. 1, p. 161-216.
- [17] GAMA, Erich, Richard Helm, Ralph Johnson, John Vlissides. *Design Patterns: Elements of Reusable Object-Oriented Software*. Addison-Wesley, 1995.
- [18] IMDB, 2012. Disponível em <<http://www.imdb.com/>>>. Acessado em <07/02/2012>.
- [19] JINNI. Find movies, TV shows matching your taste & watch online, 2012. Disponível em <<http://www.jinni.com/>>. Acessado em 05/02/2012.
- [20] MovieLens, 2012. Disponível em <<http://movielens.umn.edu/>>. Acessado em <07/02/2012>.
- [21] Whattorent, 2012. Disponível em <<http://www.whattorent.com>> Acessado em 07/02/2012.
- [22] Caraciolo, Marcel. Futuro dos Sistemas de Recomendação: Contexto e Mobilidade, 2011. Disponível em <<http://www.mobideia.com/2011/09/futuro-dos-sistemas-de-recomendacao.html>>. Acessado em 05/02/2012.
- [23] Trevisan, L. F., Cazella S. C. , Nunes, Maria A. S. N Aplicando Traços de Personalidade e Contextos em Sistemas de Recomendação para TV Digital: um facilitador do processo de ensino-aprendizagem. Disponível em <<http://www.dcomp.ufs.br/~gutanunes/hp/publications/trevisan.pdf>>. Acessado em 05/02/2012.
- [24] Dhirosattha. Predictive algorithm and what's so great about the Netflix Prize, 2009. Disponível em <<http://sendy82.wordpress.com/2009/09/22/predictive-algorithm-and-whats-so-great-about-the-netflix-prize/>>. Acessado em 07/02/2012.
- [25] Jinni. Find movies, TV shows matching your taste & watch online, 2012. Disponível em <<http://www.jinni.com/>>. Acessado em 05/02/2012.
- [26] Lawler, Nicole. THE MOVIE EXPERIENCE GETS MORE SOCIAL WITH THESMARTLIVINGROOM.COM. Disponível em <http://www.digitas.com/press/The_movie_experience_gets_more_social.pdf>. Acessado em 15/12/2012.
- [27] Likert, Rensis. A Technique for the Measurement of Attitudes, 1932. *Archives of Psychology* 140: pp. 1-55.