

PROPOSTA DE MODELO ADAPTATIVO PARA GERAÇÃO DE CONTEXTOS NA RECOMENDAÇÃO DE LOCAIS

AN ADAPTIVE PROPOSAL TO POINT-OF-INTEREST RECOMMENDER SYSTEMS

C. V. Crivelaro
celso.crivelaro@iee.org

F. J. Barth
fabricio.barth@apontador.com

R. L. A. Rocha
luis.rocha@poli.usp.br

Abstract: Recommender Systems took an important place in Internet to item recommendation for users. However, many Point-of-Interest(POI) Recommender Systems work with only one context for each user but user can be in many contexts at the same time. This work proposes an adaptive model for creation of multiples contexts for same user and thus, Recommender System can make better recommendations.

Keywords: Recommender Systems; Adaptive Technology; Geolocalization.

Resumo: Sistemas de Recomendação dominaram o mercado, especialmente na Internet para a recomendação de itens para o usuário. No entanto, muitos Sistemas de Recomendação de Locais tratam apenas de um possível contexto do usuário, mesmo este podendo estar em vários contextos ao mesmo tempo, como trabalho e viagens de férias. Esta proposta tem como objetivo propor um modelo adaptativo para a criação de múltiplos contextos para o mesmo usuário e, assim, fazer recomendações mais aderentes.

Palavras-Chave: Sistemas de Recomendação; Adaptatividade; Geolocalização.

1 INTRODUÇÃO

Os Sistemas de Recomendação ganharam grande apelo em aplicações comerciais, por recomendar novos itens aos usuários como locais para ir, produtos como livros [1], músicas e até amigos nas redes sociais.

Em muitos casos, os Sistemas de Recomendação aumentaram os acessos aos itens dos sites ou mesmo aumentaram o número de vendas como no caso do site de venda de livros Amazon [1].

As recomendações são feitas de acordo com o histórico de navegação do usuário e pela comunidade

onde o usuário está inserido. Assim, o usuário se sente familiarizado com o item oferecido. No nosso caso de estudo, serão utilizados os locais como itens para recomendação do site brasileiro Apontador.

No site Apontador, o usuário pode encontrar informações sobre locais em todo o Brasil e fazer avaliações para compartilhar a sua experiência no local com a comunidade. Um dos desafios é apresentar aos usuários os locais mais aderentes ao seu interesse de visita e que sejam mais próximos do seu cotidiano como um restaurante diferente para a hora do almoço ou um barzinho para o final de semana com os amigos ou também locais para visita durante as férias.

Ao fazer a sugestão de locais ao usuário, estamos inferindo se o usuário vai gostar do local e atribuindo um valor para ordenar os *Top N* locais que serão oferecidos. No entanto, o usuário pode modificar o seu contexto temporariamente como a verificação de locais para uma viagem ou ter dois contextos diferentes, se passa alguns dias em uma cidade e o final de semana em outra cidade. Esse fato diminui a qualidade das recomendações, pois, se a distância entre os locais sugeridos e os visitados ou avaliados for levada em consideração, pode haver recomendações fora do contexto ou foco das sugestões no contexto onde o usuário não está mais presente.

Com esse problema, esse trabalho propõe uma solução adaptativa para que o usuário possa ter vários contextos em paralelo à medida que ele visita ou avalia os locais no site e fornece mais informações sobre o seu contexto. Com mais de um contexto do usuário, é possível focar a recomendação no momento em que o usuário estiver no contexto, evitando a mistura de recomendações de locais diferentes.

Este trabalho está estruturado da seguinte maneira: na seção 2 é apresentada uma breve descrição sobre Sistemas de Recomendação; na seção 3 é fornecida uma definição de Sistema de Recomendação com contexto de locais; na seção 4 é apresentado um breve resumo sobre Tecnologia Adaptativa; na seção 5 é apresentada a proposta deste trabalho; na seção 6 são apresentadas as medidas para avaliar se essa proposta terá um ganho sobre outras propostas já conhecidas ou não.

2 SISTEMAS DE RECOMENDAÇÃO

Os Sistemas de Recomendação [2][3] surgiram no contexto de Recuperação de Informações, onde se buscavam os itens mais interessantes ao usuário baseado apenas na sua navegação e termos de busca. Esses Sistemas de Recomendação são conhecidos como Sistema de Recomendação Baseados em Conteúdo.

Nessa abordagem, são comparados os itens que o usuário interagiu com os outros itens da base. Quanto mais próximos forem esses os itens sugeridos dos itens interagidos, melhores pontuações terão para a recomendação, sendo uma abordagem interessante para recomendação de locais parecidos com o que o usuário já visitou. Para fazer essa recomendação, basta tratar os itens como documentos e fazer a similaridade entre eles baseado no modelo de espaço vetorial com pesos TF-IDF (frequência de termos – inverso da frequência do documento). [4].

Entretanto, essa abordagem tem vários problemas devido à recomendação ser feita apenas com os dados do passado do usuário e não liga outros locais que não são próximos semanticamente.

No contexto de locais, as pessoas não estão interessadas apenas em lugares iguais aos anteriores que elas visitaram, mas também em lugares em que pessoas que têm mesmo interesse vão. Para isso, há os Sistemas de Recomendação por Filtragem Colaborativa.

Com o surgimento do *Tapestry* [5], surgiu o termo Filtragem Colaborativa em que ideia é que os usuários que tiveram os mesmos interesses no passado terão os mesmos interesses no futuro. Isso implica não só em ligar os itens pelo seu conteúdo, mas ligar os itens ou os usuários pelo interesse. Com esse conceito, podemos ter recomendações de acordo com a comunidade de pessoas que vão aos mesmos lugares e oferece os lugares que estão em maior destaque no momento ao usuário que demonstra afinidade com outros usuários. Formalizando o problema, o Sistema de Recomendação faz a sugestão de lugares baseado na matriz Usuários (\mathcal{U}) e Itens (\mathcal{J}) para gerar uma lista de Recomendações (\mathcal{R}), sendo essas recomendações itens que o usuário ainda não avaliou:

$$\mathcal{U} \times \mathcal{J} \rightarrow \mathcal{R} \quad (1)$$

Um exemplo de uma matriz $\mathcal{U} \times \mathcal{J}$ pode ser mostrado abaixo:

	i_1	i_2	i_3	i_4	i_5
u_1	3	1	3	5	
u_2	4		5	4	4
u_3		3		4	3
u_4	5		4		5
u_5		2		4	

Tabela 1: Matriz Usuário X Item de avaliações

Pela Tabela 1, podemos ver as avaliações dos usuários u para os itens i . Nos espaços em branco, não houve avaliação. O objetivo do Sistema de Recomendação é inferir qual nota o usuário daria para o item não avaliado, para que ele seja recomendado ao usuário segundo o seu interesse.

Uma das formas de fazer essa inferência é usando técnicas Baseadas em Memória, como o Baseado em Vizinhança. Nessa técnica, é calculado o peso ou a similaridade entre os usuários ou itens e para produzir a recomendação é usada a média ponderada dos itens avaliados pela similaridade dos itens.

De acordo com [6], para sistemas em que existem muito mais itens do que usuários, é recomendado basear o Sistema de Recomendação em Itens, pois achar usuário correlacionados ficaria uma tarefa complexa devido a esparsidade da matriz $\mathcal{U} \times \mathcal{J}$.

Dessa forma, para se obter os locais relacionados, pode se usar cálculos de similaridade, sendo o mais comum deles a Correlação de *Pearson*:

$$w_{i,j} = \frac{\sum_{u \in \mathcal{U}} (r_{u,i} - \bar{r}_i)(r_{u,j} - \bar{r}_j)}{\sqrt{\sum_{u \in \mathcal{U}} (r_{u,i} - \bar{r}_i)^2} \sqrt{\sum_{u \in \mathcal{U}} (r_{u,j} - \bar{r}_j)^2}} \quad (2)$$

Sendo que $u \in \mathcal{U}$ é a soma de todos os usuários que coavaliaram os itens i e j , $r_{u,i}$ a avaliação que o usuário u deu ao item i e \bar{r}_i a média das avaliações para o item i .

Com um histórico de itens que o usuário já avaliou, é feita uma soma ponderada entre as similaridades do local sugerido k com as similaridades com os outros locais que o usuário já avaliou:

$$P(u, k) = \bar{r}_k + \frac{\sum_{i \in \mathcal{J}} (r_{u,i} - \bar{r}_i) w_{i,k}}{\sum_{i \in \mathcal{J}} |w_{i,k}|} \quad (3)$$

Fazendo isso para os locais relevantes, basta ordenar para que se tenha a lista dos *Top N* locais para serem sugeridos ao usuário.

Pelos itens da recomendação serem locais, o sistema de recomendação sofre o problema do contexto do usuário, ou seja, as recomendações só terão sentido se o soubermos onde o usuário está geograficamente.

Vários trabalhos fazem sugestões de locais [7][8] em que o usuário tem um dispositivo móvel que fornece a captura da sua coordenada GPS no momento de gerar a sugestão. No caso do site Apontador e de outros sites similares, muitos usuários apenas visitam as páginas dos locais, não sendo possível descobrir a sua localização exata.

3 SISTEMA DE RECOMENDAÇÃO COM CONTEXTO DE LOCAIS

Ao navegar pela busca de locais, o usuário está dando algumas dicas de como é o seu contexto ou região onde ele está, pois, irá buscar por locais onde ele deseja frequentar e irá avaliar locais onde ele frequentou e além de poder fazer *checkin* com a aplicação mobile do Apontador, indicando no momento que estava naquele local.

Diferente do contexto de aplicações como Foursquare em que sabemos onde o usuário está no momento devido aos *checkins*, a imensa maioria dos usuários do site Apontador apenas visitam as páginas dos locais nos seus computadores pessoais, o que permite visitar as páginas dos locais das mais variadas cidades e distâncias do seu local de residência. Com esse problema, é necessário inferir o contexto do usuário para que a recomendação seja correta.

Com o contexto do usuário definido, é possível inferir o quanto o local sugerido *i* pelo algoritmo colaborativo está próximo do contexto do usuário.

Como locais têm uma coordenada geográfica (X,Y) para identificação espacial, podem ser usado modelos de predição de probabilidade. Desta forma, é possível inferir se o usuário iria ao local sugerido *i* se ele visitou recentemente uma lista de locais $L = \{i_1, i_2, \dots, i_k\}$.

O modelo proposto por [9] se encaixa para essa necessidade. Para isso, foi definida uma função PL probabilidade de o usuário ir ao local *i*, pela lista de lugares L que ele já visitou ou avaliou:

$$PL(i, L) = \prod_{j \in L} P(i, j) \quad (4)$$

Sendo que P calcula a probabilidade de um usuário visitar ou avaliar os dois locais, apenas pela distância entre eles:

$$P(i, j) = a \times b^{d(i, j)} \quad (5)$$

No trabalho de [9], há uma explicação de como os parâmetros *a* e *b* são calculados para cada sistema e conjunto de dados.

A distância usada nesse modelo é a distância euclidiana já que temos as coordenadas (X,Y) dos locais *i* e *j*:

$$d(i, j) = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \quad (6)$$

Por fim, a recomendação é feita multiplicando o algoritmo colaborativo com o fator de contexto do usuário:

$$P_{final}(u, k) = P_{colaborativo}(u, k) \times P(k, L_u) \quad (7)$$

Segundo Burke [10], o Sistema de Recomendação com o contexto e o algoritmo colaborativo combinados formam um Sistema de Recomendação Híbrido com característica de cascata, ou seja, como resultado do algoritmo colaborativo é alterado pelo fator de contexto para dar o resultado final.

Esse modelo atende quando apenas analisamos o contexto cotidiano do usuário, que seriam os locais que ele vai durante o dia de trabalho e os lugares próximos à residência do usuário. Por isso, desse modelo não sabe distinguir se um novo local muito distante dos outros que o usuário visitou, é uma busca fora do contexto do usuário ou se o usuário está mudando o seu contexto, como uma mudança de cidade ou interesse em viagem de férias.

4 TECNOLOGIA ADAPTATIVA

A área de Tecnologia Adaptativa [11] tem mostrado um forte crescimento com aplicação de técnicas de aprendizado a partir de experiência anteriores e há aplicações na área de Data Mining [12].

O dispositivo adaptativo é composto de uma cada subjacente que contém um dispositivo não-adaptativo com um formalismo já conhecido que forma o núcleo do sistema e uma camada adaptativa cujas as ações modificam o núcleo apenas com as funções pré-definidas e os dados de entrada, tornando o dispositivo auto-modificável.

Assim, um dispositivo adaptativo pode ser formalizado como $DA = (ND_0, AM)$ com:

- **DA** dispositivo adaptativo
- **ND₀** dispositivo não-adaptativo subjacente;
- **AM** mecanismo adaptativo tal que $MA \text{ contido } AA \times RDN \times AP$;
- **RDN** são as regras do dispositivo não-adaptativo;
- **AA** é o conjunto das funções adaptativas anteriores e **AP** é o conjunto das funções adaptativas posteriores

As funções adaptativas contêm três regras básicas sobre o dispositivo subjacente:

1. Consulta de regra
2. Adição de regra
3. Remoção de regra

Dispositivos adaptativos são interessantes no ponto que o dispositivo subjacente pode ser usado tanto para processamento quanto para memória. O que permite que dispositivo, possa fazer tratamentos mais elaborados e se auto-modificar de acordo com o seu contexto.

Podemos ver a adaptatividade como transições no espaço de todos os possíveis dispositivo não-adaptativos em que as funções adaptativas Posterior e Anterior escolhem qual dispositivo não-adaptativo irá ser executado.

5 MODELO ADAPTATIVO PROPOSTO

Para melhorar o fator de contexto na recomendação de locais, é proposto um modelo adaptativo no dispositivo do contexto dos usuários.

A ideia é que o fator de contexto não trabalhe apenas com um contexto único como é apresentado na seção III, mas sim com vários contextos ao mesmo tempo e que esses contextos sejam inseridos durante a captura de locais que o usuário visitou ou avaliou. A adaptatividade é um conceito que se encaixa bem nessa necessidade, pois a atualização do contexto seria feita online de acordo com a navegação do usuário pelo site e a busca de locais. Dessa forma, poderíamos instantaneamente fazer recomendações com os novos contextos.

Dessa forma, cada contexto é uma lista $L_n = \{j_1, j_2, \dots, j_k\}$ com os locais que o usuário visitou, tendo tamanho máximo $\|L\| = K$, sendo K um valor inteiro arbitrário.

A definição de um contexto é que um local i dentro de L_n terá a distância euclidiana entre i e os itens j_k menor do que um valor θ arbitrário para pelo menos $\frac{\|L\|}{2}$ locais do mesmo contexto. Ou seja, todos os locais estarão dentro de uma região física. Sendo que a distância euclidiana é a distância definida entre os locais i e j pelas coordenadas (X, Y) de cada um.

Ao respeitar essa regra, não há nenhuma ação adaptativa, apenas o local que o usuário interagiu há mais tempo será retirado na lista se $\|L\| = K$, e o novo local i será adicionado. Nesse momento o dispositivo adaptativo apenas faz as ações de consulta, mas não executa nenhuma ação sobre o dispositivo subjacente que apenas inclui o novo item na sua lista, conforme apresentado na Figura 1 e Figura 2.

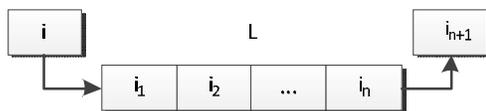


Figura 1 - Inserção de item na lista

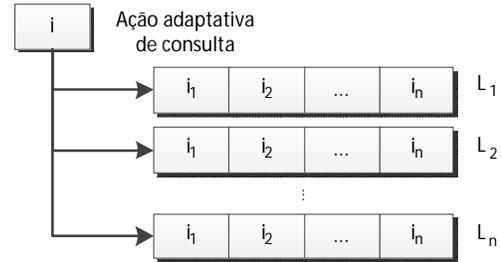


Figura 2 - Ação adaptativa de consulta

Se novo local i tiver mais do que $\frac{\|L\|}{2}$ locais distantes em um contexto, será comparado com outro contexto. Se também não encaixa dentro desse contexto, a ação adaptativa compara com outro contexto. Se em nenhum contexto o local i puder ser inserido, então um novo contexto L_{n+1} é criado.

Além de criar o novo contexto, os locais dos outros contextos tal que $D(i, j) < \theta$ serão inseridos no novo contexto, respeitando o tamanho $\|L_{n+1}\| \leq K$. Para isso serão inseridos os lugares mais próximos de i .

Nesse momento, a ação adaptativa cria uma nova regra que será o contexto L_{n+1} fazendo a inserção do local i e dos locais tal que $D(i, j) < \theta$.

Para a recomendação de um local com múltiplas listas, com o valor do algoritmo colaborativo, vamos multiplicar pelo fator de contexto. Haverá múltiplos contextos, o resultado não será apenas uma lista $TOP N$ de recomendações, mas sim n listas de recomendações.

Para o caso do sistema que tiver apenas uma lista $Top N$, apenas é necessário o contexto que tiver mais importância para aquele local i :

$$\text{Fator de Contexto} = \arg \max_n P(L_n, i) \quad (8)$$

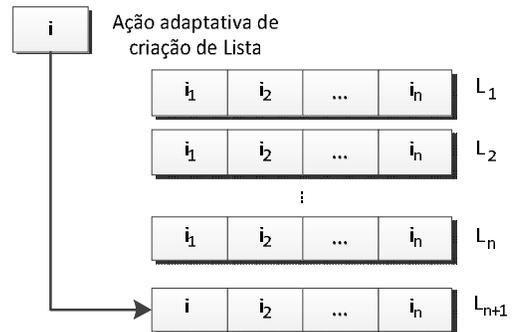


Figura 3 - Ação adaptativa de inclusão

Para fazer a recomendação final ao usuário, basta multiplicar o Fator de Contexto pelo resultado do algoritmo colaborativo, assim, os locais mais distantes

do contexto perderão importância na sugestão e os mais próximos ficarão em destaque.

Como os contextos podem ser temporários como uma viagem de férias, eles podem ser apagados depois de um tempo T . Dessa forma, se o último local inserido no contexto for inserido a mais tempo do que T , ele será eliminado. Dessa forma, evitamos a poluição de dados antigos.

6 RESULTADOS ESPERADOS

Para avaliar a melhoria da inclusão da técnica adaptativa no Sistema de Recomendação baseado em Contexto, iremos compará-lo com o mesmo Sistema de Recomendação sem o contexto para avaliar o quanto foi à melhoria.

A forma mais tradicional de medir a qualidade de um Sistema de Recomendação de acordo com [13], é medir a acurácia da predição. Uma das medidas que serão usadas é a Raiz do Erro Quadrático Médio que foi usada na avaliação do prêmio oferecido pela Netflix [14]:

$$REQM = \sqrt{\frac{1}{n} \sum_{u \in U, i \in J} (p_{u,i} - r_{u,i})^2} \quad (9)$$

Onde n é o total de avaliações sobre todos os usuários, $p_{u,i}$ é a avaliação inferida do item i ao usuário u e $r_{u,i}$ a avaliação real.

Outras medidas usadas são mais conhecidas na área de Recuperação de Informação como a Precisão e o *Recall* [15]:

- **Precisão:** Fração dos itens retornados que são relevantes
- **Recall:** Fração dos itens relevantes que são retornados

I. CONCLUSÕES

Esse trabalho propôs a inclusão de tecnologia adaptativa para a recomendação de locais. A técnica estende o conceito de contextos para que seja possível detectar mais de um contexto ao mesmo tempo em que o usuário tenha interesse.

A aplicação desse trabalho não serve apenas para a recomendação de locais, também pode ser aplicada para recomendação de outros itens como venda de produtos no e-commerce em que o usuário pode ter vários interesses distintos na compra ou pode estar comprando para outras pessoas.

REFERÊNCIAS

- [1] Linden, G., Smith, B., and York, J. Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing* 4, 1 (January 2003).
- [2] X. Su and T. M. Khoshgoftaar, A Survey of Collaborative Filtering Techniques, *Advances in Artificial Intelligence* (2009).
- [3] Adomavicius, G., Tuzhilin A.: Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions. *IEEE Transactions on Knowledge and Data Engineering*, (2005) 17(6): p. 734-749
- [4] Baeza-Yates, R. AND Ribiero-Neto, B. 1999. *Modern Information Retrieval*. Addison-Wesley Longman, Boston, Mass.
- [5] David Goldberg, David Nichols, Brian M. Oki, and Douglas Terry. 1992. Using collaborative filtering to weave an information tapestry. *Commun. ACM* 35, 12 (December 1992), 61-70.
- [6] Sarwar, B. M., Karypis, G., Konstan, J. A., AND RIEDL, J. 2001. Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th International World Wide Web Conference (WWW10)*.
- [7] C. Stiller, F. Roß, and C. Ament. Towards Spatial Awareness in Recommender Systems. In *Proceedings of the 4th International Conference for Internet Technology and Secured Transactions – ICITST'09*, Nov. 2009.
- [8] M. Brunato and R. Battiti, "Pilgrim: A Location Broker and Mobility-Aware Recommendation System", Technical Report DIT-02-092, Informatica e Telecomunicazioni, University of Trento.
- [9] M. Ye, P. Yin, W.-C. Lee, and D. L. Lee. Exploiting Geographical Influence for Collaborative Point-of-Interest Recommendation. In *Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR)*, 2011.
- [10] Burke, R., *Hybrid Recommender Systems: Survey and Experiments*, User Modeling and User-Adapted Interaction, v.12 n.4, p.331-370, November 2002
- [11] NETO, J. J., *Adaptive Rule-Driven Devices - General Formulation and Case Study*. Lecture Notes in Computer Science. Watson, B.W. and Wood, D. (Eds.): Implementation and Application of Automata 6th International Conference, CIAA 2001, Vol. 2494,
- [12] Camargo, R., Raunheite, L., *Convolução Adaptativa*. Workshop de Tecnologia Adaptativa (5 : 2011 : São Paulo) Memórias do WTA 2011. – São Paulo ; EPUSP, 2011. 111 p
- [13] Herlocker, J.L., Konstan, J.A., Terveen, L.G. and Riedl, J.T. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems*, 22 (1). 5-53
- [14] Netflix prize, <http://www.netflixprize.com/>
- [15] Ricci, F., Rokach, L., and Shapira, B., editors (2010). *Recommender Systems Handbook*. Springer.