

Simulação de um Aprendiz Informal com Recomendação Distribuída

Felipe L. Tortella¹, Carlos M. Tobar Toledo²

1. Estudante de IC da Pontifícia Universidade Católica de Campinas – PUC-Campinas; *felipetortella@gmail.com

2. Pesquisador do CEATEC da PUC-Campinas, Campinas/SP

Palavras Chave: *Modelagem de agente, simulação, recomendação distribuída.*

Introdução

O trabalho de Iniciação Científica (IC) descrito aborda uma simulação de um aprendiz informal [1], onde o mesmo se favorece de recomendações de recursos educacionais disponíveis na Web. O aprendiz informal visa aprender determinado assunto, sem estar vinculado a alguma instituição educacional. Por isso, não tem um plano de estudos definido a ser seguido. Recursos educacionais existem em grande quantidade na Web. Porém, aqueles que podem ajudar o aprendiz não são fáceis de localizar. A ideia então é que o usuário vai se beneficiar de recomendações personalizadas. A recomendação é baseada nas avaliações de outros aprendizes. Para que se faça uma recomendação efetiva, a ideia é achar aprendizes similares ao que busca recursos.

A recomendação colaborativa [2] é baseada em uma matriz de avaliações que, em geral, é esparsa devido a poucas avaliações, e muitos recursos ou usuários.

Uma forma de preencher essa matriz é através da simulação do aprendiz informal, de maneira distribuída. Considera-se então um conjunto de aprendizes simulados, cada qual em um ponto da Internet.

Para isso, cada aprendiz tem a sua matriz de recomendação, que consiste nos recursos apresentados a ele e suas avaliações a respeito desses recursos. Um aprendiz pode fazer uma avaliação de um recurso, atribuindo uma nota que vai de 1 a 10.

observam os resultados, pode-se perceber que apenas com 30 recursos essa expectativa foi cumprida. Nos outros casos, com 20 e 50 recursos, há uma pequena tendência para que isso se cumpra. Mas, com 40 e 10 recursos, os resultados não seguem padrão algum.

Isso pode ter ocorrido devido ao fato que os usuários foram selecionados aleatoriamente e que cada grupo de usuários selecionados, conforme a amostra, podia ser muito relevante para os demais usuários, gerando assim ótimas avaliações, ou ao contrário, gerando péssimas avaliações. Nos resultados com 50 usuários e 162 usuários, os resultados foram muito parecidos.

Quando se aumentou o número de usuários para 500, pôde-se perceber que as variações entre o MAE diminuem e o MAE começa a se estabilizar no valor de 1,5. Isso pode ter ocorrido porque, mesmo escondendo muitas avaliações, de recursos ainda havia um número suficiente para que se estabelecesse uma média. Isso também pôde ser observado no gráfico de 1000 usuários.

Quando o número de usuários é aumentado para 1000, pôde-se perceber que se mantém o erro no valor de aproximadamente 1,5, mas que, no caso em que os usuários se concentram em apenas 10 recursos, os valores de erro melhoram, caindo para aproximadamente 1,3. Esse fato foi observado nas outras execuções, quando se tinham mais usuários e menos recursos. Os resultados, em geral, foram melhores, pelo fato de que se tinham mais avaliações e mais usuários não tão dispersos na matriz de recomendação.

Resultados e Discussão

Os resultados aqui apresentados refletem a execução de um plano de avaliação. Dos resultados obtidos de MAE [3] (*Mean Absolute Error*) e NMAE [3] (*Normalized Mean Absolute Error*) em 125 execuções do aprendiz informal simulado é apresentado, seguido por um tipo dos gráficos obtidos, o que apresenta valores de MAE contra a quantidade de recursos, separado pela porcentagem de recursos educacionais escondidos. Foram 25 execuções para recursos em quantidade de 10, 20, 30, 40 e 50.

Para a execução do plano de avaliação foi usada uma base de dados real de recomendações [4]. A estratégia utilizada foi esconder uma determinada porcentagem de avaliações reais de cada usuário, realizar a simulação dessas avaliações, utilizando-se do algoritmo desenvolvido e confrontar os dados das avaliações reais com as simuladas. De cada 25 execuções em cinco se escondeu 30%, 40%, 50%, 60% e 70% dos recursos.

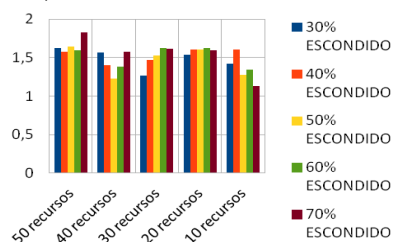


Figura 1. MAE - 100 usuários

Em uma primeira análise para as execuções envolvendo 100 usuários, Figura 1, era esperado que quanto mais avaliações escondidas, pior seria o resultado, pois, quando o sistema fosse gerar as novas avaliações de recursos, existiriam menos avaliações para comparação. Quando se

Conclusões

O trabalho possibilitou um entendimento do que são recomendadores, recursos educacionais e a concepção da ideia do aprendiz informal e sua modelagem.

Outros fatores que também estavam entre os objetivos, foram a simulação do aprendiz informal e a confrontação dos resultados dessa simulação com dados reais.

Os resultados da avaliação do sistema não corresponderam ao que se esperava. O grande motivo deve ser a quantidade de avaliações de recursos, que variou a cada execução, por conta da escolha aleatória dos usuários a serem considerados. Notou-se também a necessidade de bases de recursos com mais tipos de dados, que cubram aqueles projetados no sistema.

Agradecimentos

Agradeço ao CNPq pela bolsa oferecida, à PUC-Campinas pela oportunidade e ao meu orientador que me proporcionou, além do conhecimento obtido durante a IC, um grande despertar pela pesquisa.

- [1] Manouselis, N.; Drachler, H.; Vuorikari, R. Hummel, H. G. K.; Koper, R., (2011), *Recommender System in Technology Enhanced Learning*, In P. B. Kantor, F. Ricci, L. Rocach, B. Shapira (Eds.), *Recommender System Handbook*, p. 386-415, Berlin: Springer.
- [2] Zhang, Y; Liu, Haiou; Li, Shiyong (2011), *A Distributed Collaborative Filtering Recommendation: Mechanism for Mobile Commerce Based on Cloud Computing*, *Journal of Information & Computational Science* 8.
- [3] Goldberg, K, (2001), *Eigentaste: A Constant Time Collaborative Filtering Algorithm*, *Information Retrieval*, p. 133-15.
- [4] Merlot. Disponível em: <<http://www.merlot.org>>. Acesso em: 13 dez. 2013.