Uma Metodologia de Recomendação Híbrida Para Sistemas de Armazenamento em Nuvem

Alternative title: An Hybrid Methodology Recommendation in Cloud Storage Systems

Ricardo Batista Rodrigues
Centro de Informática - Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
rbr@cin.ufpe.br

Vinicius Cardoso Garcia
Centro de Informática - Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
vcg@cin.ufpe.br

Silvio R. de Lemos Meira
Centro de Informática - Universidade Federal de Pernambuco (UFPE)
Recife, PE, Brasil
srlm@cin.ufpe.br

ABSTRACT

The filtering relevant content amid a big database is a complex, hard-working and time consuming. With the spread of the use of cloud storage systems, big databases migrated to these systems and the data also followed the problems encountered in filtering relevant content in big data sets. This paper proposes a hybrid methodology for recommendation files in cloud storage systems using the main recommendation techniques and cloud features.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Systems; H.3 [**Information Storage and Retrieval**]: Miscellaneous

General Terms

Theory

Keywords

Recommendation System, cloud storage, cloud computing.

RESUMO

A filtragem de conteúdo relevante em meio uma grande base de dados é complexa, trabalhosa e demanda tempo. Com a propagação do uso de sistemas de armazenamento em nuvem, grandes volumes de dados foram migrados para esses sistemas, e com os dados também seguiram os problemas encontrados na filtragem de conteúdo relevante em grandes massas de dados. Este trabalho apresenta uma proposta de metodologia para recomendação de arquivos em sistemas de armazenamento em nuvem, utilizando as principais técnicas de recomendação e características da nuvem.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
.

Palavras-Chave

Sistemas de recomendação, armazenamento em nuvem, computação em nuvem.

1. INTRODUÇÃO

Com a propagação do uso de sistemas de armazenamento em nuvem, grandes volumes de dados foram migrados para esses sistemas e com os dados também seguiram os problemas. Alguns dos problemas encontrados em grades volumes de dados em sistemas de armazenamento em nuvem são: filtragem de conteúdo relevante, disponibilidade, elasticidade e segurança. Um estudo realizado pela KPMG International [1] apontou que cerca de 50% das empresas participantes opta por armazenar dados em nuvem. Com essa crescente utilização de armazenamento em nuvem, são formados grandes volumes de dados.

A escassez de informação disponível deu lugar a uma imensa massa de dados ao alcance de todos. Entretanto, esta inversão também acabou gerando um problema: como encontrar dados relevantes em meios a uma grande massa de dados? A filtragem de conteúdo relevante em meio a essa imensidão de dados é complexa, demanda tempo e recursos computacionais. Outro ponto a ser considerado neste cenário é como proporcionar aos usuários uma melhor utilização dos recursos em nuvem, na enfadonha tarefa de filtrar conteúdo relevante em meio a uma imensidão de dados [10].

Um forma de amenizar os problemas com a filtragem de conteúdo em um grande conjunto de dados, pode ser o uso de técnicas de recomendação. Para isto, é necessário informações sobre o indivíduo alvo da recomendação ou sobre o ambiente que influenciará na geração da recomendação. A partir destas informações, um sistema de recomendação (SR) poderá recomendar arquivos que apresentem maior semelhança com as preferências do usuário alvo [8].

Este trabalho apresenta uma proposta de pesquisa que tem como objetivo propor uma metodologia de recomendação de arquivos para sistemas de armazenamento em nuvem, utilizando as principais técnicas de recomendação e características de ambientes de armazenamento em nuvem. O objetivo deste modelo é recomendar aos usuários arquivos

[1] https://kpmg.com

que sejam similares a suas preferências e em paralelo proporcionar uma melhor utilização dos recursos do ambiente em nuvem.

Este artigo está organizado da seguinte forma: a Seção 2 apresenta a fundamentação teórica. Na Seção 3 é apresentado o problema de pesquisa. Na Seção 4 é apresentada a proposta de pesquisa. Na Seção 5 é apresentada a proposta de avaliação da pesquisa. Na seção 6 são apresentadas as principais contribuições. Na Seção 7 é apresentado o estado atual do trabalho. A Seção 8 apresenta os trabalhos relacionados. Na Seção 9 são apresentadas as considerações finais.

2. FUNDAMENTAÇÃO TEÓRICA

Nesta Seção, são apresentados os principais conceitos desta trabalho: computação em nuvem, sistemas de armazenamento em nuvem e as principais técnicas utilizadas em sistemas de recomendação.

2.1 Computação em Nuvem

O National Institute of Standards and Technology (NIST)[2] define computação em nuvem como um modelo que permite que um conjunto de recursos computacionais possam ser fornecidos sob demanda de forma a permitir que os mesmos sejam fornecidos e liberados rapidamente com o mínimo de esforço de gestão ou interação do fornecedor [11]. Vaquero et al.[16] define computação em nuvem como um grande conjunto de recursos virtualizados (*hardware*, plataformas de desenvolvimento e/ou serviços) facilmente usáveis e acessíveis. A arquitetura da computação em nuvem é comumente representada em camadas: *Software* como serviço, Plataforma como Serviço e Infraestrutura como serviço.

- **Iaas:** É o fornecimento de recursos computacionais como: capacidade de processamento, armazenamento e conectividade. Tomando armazenamento como um exemplo, quando o usuário usa um serviço de armazenamento em computação em nuvem, ele somente paga a parte que foi efetivamente consumida sem comprar nenhum disco rígido ou até mesmo sem nem conhecer a localização dos seus dados armazenados. Um exemplo de IaaS é o Amazon Web Services (Amazon AWS)[3], principalmente, através do Amazon Elastic Compute Cloud (Amazon EC2) [4];

- **Paas:** Fornece a infraestrutura e dá suporte a um conjunto de interfaces de programação para aplicações em nuvem. É o meio entre o hardware e as aplicações. Um bom exemplo de PaaS é a plataforma de serviços do Microsoft Azure [5];

- **Saas:** Em SaaS o objetivo é substituir as aplicações que rodam nos computadores. Ao invés de comprar o software por um preço relativamente alto, o usuário paga pelo que é usado do SaaS, o que pode gerar economia, observando que o usuário só pagará pelo que usar. Um dos pincipais problemas é a baixa taxa de transmissão de dados na rede em determinadas localidades o que é fatal para algumas aplicações, como sistemas de tempo real. Um exemplo muito conhecido de SaaS é o Dropbox [6];

[2]http://www.nist.gov/srm/
[3]http://aws.amazon.com/
[4]http://aws.amazon.com/ec2/
[5]http://www.windowsazure.com/

2.2 Sistemas de Armazenamento em Nuvem

Entre os recursos disponiveis na técnologia em nuvem, está o de armazenamento, o qual provê recursos e serviços de armazenamento baseados em servidores remotos que utilizam os princípios da computação em nuvem [18]. Armazenamento em nuvem tem duas características básicas: a primeira trata da infraestrutura da nuvem, a qual baseia-se em clusters de servidores baratos; a segunda tem o objetivo de, através dos *clusters* de servidores, armazenamento distribuído e redundância de dados, fazer múltiplas cópias dos dados armazenados para alcançar dois requisitos: alta escalabilidade e alta usabilidade. A alta escalabilidade significa que o armazenamento em nuvem pode ser dimensionado para um grande aglomerado com centenas de nós ou peers de processamento. Alta usabilidade significa que o armazenamento em nuvem pode tolerar falhas de nós e que estas falhas não afetam todo o sistema [5].

Com a expansão do mercado e das tecnologias em nuvem surgiram diversos sistemas de armazenamento de dados em nuvem, como, por exemplo, DropBox[7], Box[8], Mozy[9], JustCloud[10] e Ustore[11], estes são alguns dentre os diversos sistemas de armazenamento em nuvem existentes.

2.3 Sistemas de Recomendação

Sistemas de Recomendação (SR) são ferramentas e técnicas de software que fornecem recomendação de itens. As sugestões geradas por um SR podem estar relacionadas a processo de tomada de decisão em diversos contextos, como, a escolha de itens em *e-commerces*, conteúdo similar as preferência do usuário, livros, amigos em redes sociais dentro outras opções [13] .

Uma recomendação pode se basear nas preferências de quem a faz e pode ser dirigida a um indivíduo específico, ou para um público maior. Para a pessoa que recebe a recomendação, ela funciona como um filtro ou uma visão particular de um universo de possibilidades geralmente inacessível. Ela pode levar em consideração também a preferência de quem está à procura de sugestões e não apenas de quem a faz. É possível até mesmo fazer recomendação baseada nas opiniões de outras pessoas. Alguém que não é admirador do gênero Rock pode recomendar discos baseado no que seus amigos que apreciem tal estilo costumam ouvir. Ainda, a recomendação pode incluir explicações sobre como ela foi gerada para permitir que o seu recebedor a avalie [13].

No universo dos sistemas de recomendação, existem diversas técnicas e formas de gerar recomendação, dentre quais destacam-se: recomendação baseada em conteúdo, filtragem colaborativa e filtragem híbrida [3] [13].

Filtragem baseada em conteúdo: são sistemas que fazem recomendações ao usuário de itens semelhantes àqueles em que ele demonstrou interesse no passado. Para tanto, o sistema analisa as descrições dos conteúdos dos itens avaliados pelo usuário para montar o seu perfil, o qual é utilizado para filtrar os demais itens da base. Esse conteúdo no qual

[6]http://www.dropbox.com
[7]https://www.dropbox.com/
[8]https://www.box.com/
[9]http://www.mozy.com/
[10]http://www.justcloud.com/
[11]http://usto.re/

ele se baseia são elementos explícitos como nome, descrição, *tags*, conteúdo, categorização ou *rating* do item a ser recomendado. Os resultados são o julgamento da relevância daqueles itens para o usuário, e a consequente recomendação ou não [12] [13].

Uma das vantagens de implantar um filtro baseado em conteúdo é que a quantidade de usuários no sistema não interfere na eficácia do SR, já que se baseia somente no histórico do que o usuário já acessou. Em contrapartida, um sistema assim precisa de itens bem descritos, com informação suficiente para categorizá-los. Outro problema encontrado nesse tipo de recomendação é a sugestão de itens sempre muito parecidos, limitando os usuários de conhecer itens diferentes [13].

Filtragem colaborativa: a recomendação por filtragem colaborativa foi proposta inicialmente para suprir as deficiências da abordagem baseada em conteúdo. Com o passar dos anos, conquistou tamanha aceitação que hoje é provavelmente a técnica mais amplamente conhecida, implementada e utilizada para sistemas de recomendação [1] [6].

Na abordagem colaborativa, em contraste com a recomendação baseada em conteúdo, a compreensão ou conhecimento do conteúdo dos itens é totalmente prescindível. Ao invés de buscar itens disponíveis com conteúdos similares aos previamente avaliados positivamente pelo usuário para indicá-los, ela se apoia inteiramente na similaridade entre os usuários do sistema para o processo de sugestão. Partindo do princípio de que as melhores recomendações para um indivíduo são aquelas feitas por pessoas com preferências similares às dele, o sistema identifica estas pessoas para sugerir itens que as mesmas tenham aprovado e ainda não tenham sido consumidos pelo indivíduo [13].

Filtragem híbrida: são sistemas de recomendação que utilizam duas ou mais técnicas, para amenizar os problemas apresentados por cada técnica [13].

3. APRESENTAÇÃO DO PROBLEMA

Diversas aplicações buscam proporcionar serviços com o objetivo de facilitar atividades regulares e proporcionar conforto aos seus usuários. Este cenário se confirma com a expansão da vida social na Internet e cada vez mais a utilização de sistemas em nuvem. Os sistemas de armazenamento em nuvem, além de prover diversas vantagens em relação a outras formas de armazenar arquivos, devem seguir este cenário na filtragem por conteúdo relevante em meio a grades volumes de arquivos.

Como forma de prover facilidades e conforto ao usuário no uso de sistemas, destacam-se os sistemas de recomendações. Um SR tem como objetivo fornecer sugestões de itens que possam ser similares aos interesses do usuário [4] [13].

A maioria dos sistemas de armazenamento em nuvem, como, Amazon S3[12], Dropbox[13], SugarSync[14], entre outros, não oferecem serviços de recomendação de arquivos ao usuário. A maioria somente permite que o usuário realizem consultas por um determinado arquivo a partir de palavras-chave. Isso torna a busca por conteúdo relevante, repetitiva e trabalhosa, demandando certa quantidade de tempo na filtragem de conteúdo. Em algumas buscas o arquivo encontrado pelo usuário pode não proporcionar a melhor utilização dos recursos do sistema e de seus recursos computacionais, como, por exemplo, o *download* do arquivo, que pode demandar mais tempo e recursos do que o disponível pelo usuário [14] [15].

[12]http://aws.amazon.com/pt/s3/
[13]https://www.dropbox.com/
[14]https://www.sugarsync.com/

4. PROPOSTA DE SOLUÇÃO

A proposta de pesquisa deste trabalho tem como objetivo principal apresentar uma metodologia para recomendação de arquivos em sistemas de armazenamento em nuvem. A metodologia será composta por características que representem as duas principais técnicas utilizadas em mecanismos de recomendação (Filtragem baseada em conteúdo e Filtragem Colaborativa), associadas a características que representem o ambiente de armazenamento em nuvem. Para atingir os objetivos propostos, esta pesquisa foi divida em cinco etapas:

Primeira etapa - Revisão sistemática da literatura: nesta etapa, foi desenvolvido um estudo sistemática da literatura. O objetivo foi identificar modelos e mecanismos de recomendação que utilizem características de ambientes em nuvem. A partir da execução desta etapa, foram definidas características da nuvem para compor a metodologia de recomendação. As características resultantes desta etapa foram: taxa de disponibilidade, taxa de *download* e o tamanho dos arquivos.

Segunda etapa - Modelo baseado em conteúdo e em nuvem: nesta etapa, foi proposto um modelo de recomendação baseado em conteúdo e em caraterísticas do ambiente em nuvem. O modelo proposto nesta etapa foi denominado RecCloud, composto pelas características da primeira etapa e características que representam a técnica de recomendação baseada em conteúdo.

Terceira etapa - Modelo baseado em filtragem colaborativa e nuvem: nesta etapa, será proposto um modelo de recomendação baseado em filtragem colaborativa e em características do ambiente em nuvem.

Quarta etapa - Avaliação dos resultados: nesta etapa, serão avaliados os resultados obtidos nas etapas anteriores. O objetivo é identificar quais critérios de recomendação utilizados nos modelos propostos obtiveram e colaboraram para os melhores resultados. Nesta etapa, será realizado um estudo sistemático com objetivo de identificar as principais formas de avaliação utilizadas em sistemas de recomendação. A partir das técnicas encontradas nesta etapa, será definida a forma de avaliação que será adota nesta pesquisa.

Quinta etapa - Metodologia híbrida: nesta etapa, será proposta uma metodologia híbrida, composta pelas características e critérios utilizados nos modelos desenvolvidos nas etapas anteriores.

5. PROJETO DE AVALIAÇÃO

Para avaliação deste trabalho será utilizada a metodologia proposta por Jain [7], onde é defendido que para realizar uma avaliação é preciso ter objetivos, métricas, fatores e níveis.

Na literatura são encontradas diversas técnicas para avaliação de mecanismos de recomendação, algumas das mais conhecidas são: Precisão, *Recall*, *F-measure*, Curva Roc, métrica de Breese, Npdm e o feedback do destinatário alvo da recomendação [17]. Estas são algumas das opções dentre uma vasta lista de técnicas para avaliar recomendações.

Neste trabalho, propomos a utilização das métricas de

Precisão, *Recall*, *F-measure* e avaliar o tempo gasto no *download* dos artigos recomendados. Estas técnicas, foram escolhidas por serem as que mais se adequaram a metodologia proposta, por analisarem a precisão de cada *ranking* de recomendação gerado, o tempo amenizado no *download* dos arquivos recomendados, além de avaliar a eficiência do modelo. Outras métricas de avaliação podem ser adotadas no decorrer desta pesquisa. A seguir são apresentadas as técnicas citadas:

Precisão: é a taxa de itens relevantes recomendados no resultado. É dada, através da proporção entre o número de arquivos relevantes recomendados e o número total de arquivos recomendados [2] [9]. O cálculo da métrica Precisão é apresentado na Equação 1

$$Precisao = \frac{\mid \{arqrelevantes\} \bigcap \{arqrecomendados\} \mid}{\{arqrecomendados\}} \quad (1)$$

Na Equação 1, "*arqrelevantes*" é a quantidade de arquivos recomendados que fazem parte do ranque de relevância, e "*arqrecomendados*" é a quantidade de arquivos recomendados para cada solicitação de recomendação. O resultado, é representado por valores entre 0 (zero) e 1 (um), quanto mais próximo de 1 (um), mais preciso é o sistema.

***Recall*:** é a taxa de itens relevantes recomendados em relação à quantidade total de itens relevantes [9]. O cálculo da métrica *Recall* é apresentado na Equação 2.

$$Recall = \frac{\mid \{arqrelevantes\} \bigcap \{arqrecomendados\} \mid}{\{arqrelevantes\}} \quad (2)$$

Na Equação 2, "*arqrelevantes*" é a quantidade de arquivos recomendados que fazem parte do ranque de relevância, e "*arqrecomendados*" é a quantidade de arquivos recomendados para cada solicitação. O resultado é representado por valores entre 0 (zero) e 1 (um), quanto mais próximo de 1 (um) mais o sistema satisfaz a solicitação da recomendação.

***F-measure*:** é a média ponderada das taxas de Precisão e *Recall*. O cáculo da métrica *F-measure* é apresentado na Equanção 3

$$F - measure\ \alpha = \frac{(1+\alpha) \cdot precisao \cdot recall}{(\alpha \cdot precisao) + recall} \quad (3)$$

Tempo gasto no *download*: serão efetuados *downloads* dos arquivos recomendados e observado o tempo gasto para realizar cada *download*. Os resultados obtidos serão comparados com o tempo gasto no *download* de arquivos recomendados utilizando todos os modelos propostos nesta pesquisa.

6. CONTRIBUIÇÕES

No desenvolver desta pesquisa será apresentado como contribuição um estudo sistemático sobre a utilização de sistemas de recomendação em nuvem, com ênfase em sistemas de armazenamento e em técnicas de avaliação de modelos de recomendação.

Nesta pesquisa, será apresentado como contribuições dois modelo de recomendação. Um modelo será composto pela técnica de recomendação baseada em conteúdo associada a características do ambiente em nuvem. O segundo modelo será baseado em filtragem colaborativa associada às características do ambiente em nuvem.

Ao final desta pesquisa, é esperado como resultado final uma metodologia para recomendação de arquivos em sistemas de armazenamento em nuvem. A metodologia esperada deverá ser composta pelas características utilizadas nos dois modelos apresentados no decorrer da pesquisa. Desta forma, será composta pelas técnicas de recomendação baseada em conteúdo e recomendação baseada em filtragem colaborativa associadas às características de sistemas de armazenamento em nuvem.

7. ATIVIDADES REALIZADAS

A pesquisa encontra-se em desenvolvimento, a primeira etapa deste trabalho consistiu na revisão da literatura de forma sistemática, com o objetivo de melhor entender os conceitos de recomendação e computação em nuvem. O estudo sitemático auxiliou a fundamentação da proposta de pesquisa, mapear trabalhos relacionados, e montar o referencial teórico desta pesquisa. A revisão da literatura foi realizada em um processo de ampla pesquisa à procura de trabalhos publicados, combinando pesquisa automática e manual. A busca automática foi realizada nas fontes de pesquisa: IEEEXplore Digital Library[15], ACM Digital Library[16], Science Direct[17], Scopus[18]. A busca manual foi realizada nos anais das conferências: *The ACM Conference Series on Recommender Systems - RecSys* [19] e *IEEE International Conference on Cloud Computing*[20], por serem conferências de destaque nas áreas desta pesquisa.

A partir da revisão da literatura, foram encontradas algumas características de sistemas de armazenamento em nuvem que podem ser utilizadas como métricas em mecanismos de recomendação. Dentre as encontradas, três foram utilizadas como proposta para representarem os sistemas de armazenamento em nuvem, as características selecionadas foram as mais citadas nos trabalhos encontrados. A seguir são apresentadas as características escolhidas:

Disponibilidade: refere-se ao tempo em que um arquivo estará disponível para o usuário. A disponibilidade pode ser medida em horas, ou seja, a quantidade de horas em que um arquivo a ser recomendado está disponível na nuvem. Um arquivo somente deve ser recomendado ao usuário se estiver disponível e possibilitando o seu *download*. O critério disponibilidade representa uma das principais características e anseios quanto à tecnologia de computação em nuvem. A maioria dos usuários que migram para a nuvem são atraídos pela oferta de altas taxas de disponibilidade, elasticidade e mobilidade, o que torna possível armazenar arquivos em grande quantidade e que estejam disponíveis e acessíveis a qualquer momento a partir de uma conexão com a Internet [11]. **Taxa de *Download*:** o objetivo desta característica é que arquivos que proporcionam melhores condições para a redução no tempo gasto no *download* fiquem melhor ranqueados que os demais arquivos. A contribuição deste critério na redução do tempo gasto no *download* de um arquivo recomendado, é produzida em conjunto com o critério "Tamanho do Arquivo", apresentado no próximo item. **Tamanho do Arquivo:** esta característica tem como objetivo contribuir

[15] http://ieeexplore.ieee.org/Xplore/home.jsp
[16] http://dl.acm.org/
[17] http://www.sciencedirect.com/
[18] http://www.scopus.com/
[19] http://recsys.acm.org/
[20] http://www.thecloudcomputing.org/2015/

na tarefa de amenizar o tempo gasto no *download* de um arquivo recomendado. Como explicado no item critério "Taxa de Download", os dois critérios estão diretamente relacionados. O ranque de recomendação mudará de acordo com à taxa disponível para *download* e o tamanho dos arquivos candidatos a recomendação.

A segunda etapa realizada consistiu na proposta de um modelo de recomendação baseado em conteúdo e em nuvem. O modelo proposto é composto por cinco critérios de recomendação, os três propostos na primeira etapa e mais dois para representar a técnica de recomendação baseada em conteúdo. Os dois critérios são:

- Similaridade: este critério atende ao requisito referente as preferências do usuário. Neste critério é calculada a similaridade entre o conteúdo de um arquivo no qual o usuário tenha demostrado preferência e arquivos candidatos a serem recomendados.
- Popularidade do arquivo: este critério representa a importância social de um arquivo na nuvem, avaliado por meio da quantidade de *downloads* que foram realizados de um mesmo arquivo.

Para a avaliação do modelo, foi implementado um protótipo. Foram implementados os critérios propostos no sistema de armazenamento em nuvem Ustore[21].

Na avaliação preliminar do trabalho, foram utilizadas as técnicas: Precisão, *Recall* e *F-measure*. Os melhores resultados obtidos foram: Precisão = 0.68; Recall = 0.42, F-measure = 0.42.

O pesquisa encontra-se em sua terceira etapa de desenvolvimento. Nesta etapa, está sendo realizado um estudo sistemático sobre sistemas de recomendação baseados em filtragem colaborativa, em seguida será proposto, implementado e avaliado um modelo de recomendação baseado em filtragem colaborativa.

8. TRABALHOS RELACIONADOS

Existem diversos sistemas de recomendação disponíveis na Internet, e boa parte destes sistemas estão relacionados à nuvem, seja como parte de sistemas em nuvem ou hospedados na nuvem. Muitos deste utilizam dados da nuvem em suas recomendações ou características da nuvem para gerar recomendações. A proposta desta pesquisa se diferência utilizar características da nuvem na geração de suas recomendações. Desta forma, cada critério é parte da recomendação, que tem como objetivo proporcionar aos usuários a melhor utilização dos recursos em nuvem disponíveis.

9. CONSIDERAÇÕES FINAIS

Este artigo apresentou uma proposta de pesquisa que tem como objetivo propor uma metodologia de recomendação para sistemas de armazenamento em nuvem. Apresentamos as técnicas (Recomendação baseada em conteúdo e filtragem colaborativa) que serão utilizadas no desenvolvimento da metodologia proposta, essas técnicas serão associadas a características de sistemas de armazenamento em nuvem. O objetivo principal é proporcionar uma melhora utilização do ambiente em nuvem e amenizar o tempo gasto na filtragem de conteúdo.

[21] http://usto.re/

10. REFERÊNCIAS

[1] A. Ansari, S. Essegaier, and R. Kohli. Internet recommendation systems. *JOURNAL OF MARKETING RESEARCH*, 37(3):363–375, 2000.

[2] R. A. Baeza-Yates and B. Ribeiro-Neto. *Modern Information Retrieval*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

[3] R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, 2002.

[4] S. C. Cazella, E. Reategui, and L. O. C. Alvares. E-commerce recommender's authority: applying the user's opinion relevance in recommender systems. In *Proceedings of the 12th Brazilian Symposium on Multimedia and the web*, WebMedia '06, pages 71–78, New York, NY, USA, 2006. ACM.

[5] J. Deng, J. Hu, A. Liu, and J. Wu. Research and application of cloud storage. In *Intelligent Systems and Applications (ISA), 2010 2nd International Workshop on*, pages 1–5, 2010.

[6] J. L. Herlocker. *Understanding and improving automated collaborative filtering systems*. PhD thesis, University of Minnesota, 2000. AAI9983577.

[7] R. K. Jain. *The Art of Computer Systems Performance Analysis: Techniques for Experimental Design, Measurement, Simulation, and Modeling*. Wiley, 1 edition, Apr. 1991.

[8] G. Jung, T. Mukherjee, S. Kunde, H. Kim, N. Sharma, and F. Goetz. Cloudadvisor: A recommendation-as-a-service platform for cloud configuration and pricing. In *Services (SERVICES), 203 IEEE Ninth World Congress on*, pages 456–463, 2013.

[9] S. Lee, D. Lee, and S. Lee. Personalized dtv program recommendation system under a cloud computing environment. *IEEE Trans. on Consum. Electron.*, 56(2):1034–1042, May 2010.

[10] A. R. S. Lopes. Sistemas de recomendaï£¡ao utilizando similaridade globais para aliviar o problema da esparcidade. Master's thesis, Universidade Federal de Pernambuco (UFPE), 2012.

[11] P. Mell and T. Grance. The NIST Definition of Cloud Computing. Technical report, National Institute of Standards and Technology, Information Technology Laboratory, July 2009.

[12] M. J. Pazzani and D. Billsus. Learning and revising user profiles: The identification of interesting web sites. *Machine Learning*, 27(3):313–331, 1997.

[13] F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors. *Recommender Systems Handbook*. Springer, 2011.

[14] R. B. Rodrigues, C. M. R. da Silva, F. A. Durao, R. E. Assad, V. C. Garcia, and S. R. L. Meira. A file recommendation model for cloud storage systems. In *Proceedings of the Annual Conference on Brazilian Symposium on Information Systems: Information Systems: A Computer Socio-Technical Perspective - Volume 1*, SBSI 2015, pages 16:111–16:118, Porto Alegre, Brazil, Brazil, 2015. Brazilian Computer Society.

[15] R. B. Rodrigues, F. A. Durao, V. C. Garcia, C. M. R. Silva, R. R. Souza, and R. E. Assad. A cloud-based

recommendation model. In *Proceedings of the 7th Euro American Conference on Telematics and Information Systems*, EATIS '14, pages 23:1–23:4, New York, NY, USA, 2014. ACM.

[16] L. M. Vaquero, L. Rodero-Merino, J. Caceres, and M. Lindner. A break in the clouds: towards a cloud definition. *SIGCOMM Comput. Commun. Rev.*, 39(1):50–55, Dec. 2008.

[17] Y. Y. Yao. Measuring retrieval effectiveness based on user preference of documents. *J. Am. Soc. Inf. Sci.*, 46(2):133–145, Mar. 1995.

[18] W. Zeng, Y. Zhao, K. Ou, and W. Song. Research on cloud storage architecture and key technologies. In *Proceedings of the 2Nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, ICIS '09, pages 1044–1048, New York, NY, USA, 2009. ACM.