

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

GUILHERME HAAG RIBACKI

**Uma Abordagem de Recomendação de
Colaborações Acadêmicas Através da
Análise de Séries Temporais**

Dissertação apresentada como requisito
parcial para a obtenção do grau de Mestre em
Ciência da Computação

Orientador: Prof. Dr. Leandro Krug Wives

Porto Alegre
2016

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Haag Ribacki, Guilherme

Uma Abordagem de Recomendação de Colaborações Acadêmicas Através da Análise de Séries Temporais / Guilherme Haag Ribacki. – Porto Alegre: PPGC da UFRGS, 2016.

60 f.: il.

Dissertação (mestrado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação, Porto Alegre, BR–RS, 2016. Orientador: Leandro Krug Wives.

1. Sistemas de Recomendação. 2. Recomendação de Colaborações Acadêmicas. 3. Redes Sociais. 4. Redes de Colaboração. 5. Séries Temporais. 6. Sistemas de Recomendação Sensíveis ao Tempo. I. Wives, Leandro Krug. II. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos Alexandre Netto

Vice-Reitor: Prof. Rui Vicente Oppermann

Pró-Reitor de Pós-Graduação: Prof. Vladimir Pinheiro do Nascimento

Diretor do Instituto de Informática: Prof. Luis da Cunha Lamb

Coordenador do PPGC: Prof. Luigi Carro

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

AGRADECIMENTOS

Agradeço, primeiramente, à minha família, que sempre me apoiou e me incentivou durante toda a carreira acadêmica. Em especial aos meus pais, à minha tia e segunda mãe, ao meu avô e tutor, e aos meus dois irmãos. Graças a eles sou a pessoa que sou hoje. Agradeço também aos outros membros da minha família pelo carinho e incentivo para estudar e perseguir meus objetivos.

Agradeço especialmente à minha esposa pelo apoio, carinho, compreensão, incentivo e companheirismo, e por ter estado presente durante praticamente toda a minha jornada acadêmica. Eu não teria chegado até aqui sem ela.

Agradeço aos meus amigos, por toda a compreensão e apoio, que mesmo frente à ausência por longos períodos jamais deixaram de me apoiar.

Agradeço aos colegas de graduação e mestrado, que sempre tiveram algo a contribuir e me ensinar, pelas colaborações nos trabalhos das disciplinas e parceria fora do contexto acadêmico.

Agradeço aos professores e mestres que encontrei durante a minha jornada acadêmica, pelo conhecimento transmitido e pelas oportunidades de aprendizado. Todos eles tiveram um papel importante na minha formação.

Agradeço à CAPES pelos dois meses de bolsa de estudos durante o curso de mestrado.

Por último, gostaria de agradecer ao meu professor orientador, por todo apoio e incentivo prestado desde os últimos semestres da graduação até o presente momento. Obrigado!

RESUMO

O avanço da tecnologia nos últimos anos permitiu a criação de Sistemas de Informação com acesso a grandes bases de dados, abrindo diversas possibilidades de aplicações. Tem-se como exemplo a Internet, onde uma enorme quantidade de dados é gerada e publicada a todo momento por usuários ao redor do mundo. Com isso, aos poucos foi surgindo a necessidade de métodos para filtrar o conteúdo disponível de forma a permitir que um usuário pudesse focar apenas nos seus interesses. Nesse contexto surgiram os Sistemas de Recomendação e as Redes Sociais, onde, mais recentemente, surgiram trabalhos que apresentam abordagens para o uso de Sistemas de Recomendação no contexto acadêmico, de forma a aumentar a produtividade de grupos de pesquisa. Também têm sido bastante exploradas formas de se utilizar informações temporais em Sistemas de Recomendação de maneira a melhorar as recomendações feitas. O presente trabalho propõe uma abordagem de recomendação de colaborações acadêmicas utilizando a técnica de Análise de Séries Temporais, buscando melhorar os resultados obtidos por trabalhos anteriores. Foi realizado um experimento *offline* para avaliar o desempenho da abordagem proposta em relação às abordagens anteriores e um estudo de usuários para fazer uma análise mais profunda com *feedback* de usuários. Foram utilizadas métricas conhecidas das áreas de Recuperação de Informação e Sistemas de Recomendação, mas alguns resultados se mostraram inferiores em comparação com as abordagens existentes; outros, porém, foram similares. Também foram utilizadas algumas métricas de avaliação focadas em Sistemas de Recomendação, e os resultados obtidos foram similares em todas as abordagens testadas.

Palavras-chave: Sistemas de Recomendação. Recomendação de Colaborações Acadêmicas. Redes Sociais. Redes de Colaboração. Séries Temporais. Sistemas de Recomendação Sensíveis ao Tempo.

An Approach For Academic Collaborations Recommendation Through Time-Series Analysis

ABSTRACT

The advance of technology in recent years made possible the creation of Information Systems with access to large databases, opening many applications possibilities. There's the Internet, for example, where a vast amount of data is generated and published all the time by users around the world. In this sense, the need for methods to filter the available content to enable users to focus only on their interests slowly emerged. In this context, Recommender Systems and Social Networks appeared, where, recently, works reporting approaches to provide recommendations in the academic context appeared, increasing the productivity of research groups. New ways to employ temporal information in Recommender Systems to make better recommendations are also being explored. The present work proposes an approach to academic collaborations recommendation using Time Series Analysis, aiming to improve results reported on previous and current works. An offline experiment was done to evaluate the proposed approach in comparison with other works and a user study was done to make a deeper analysis from user feedback. Known metrics from the Information Retrieval and Recommender Systems fields were used, and in some cases the results obtained were lower compared to the current methods but similar in others. Some evaluation metrics from Recommender Systems were also used, and the results were similar to all approaches.

Keywords: Recommender Systems, Academic Collaboration Recommendation, Social Networks, Collaboration Networks, Time-Series, Time-Aware Recommender Systems.

LISTA DE ABREVIATURAS E SIGLAS

AR	<i>Autoregressive</i>
MA	<i>Moving Average</i>
ARMA	<i>Autoregressive moving average</i>
ARIMA	<i>Autoregressive integrated moving average</i>
CORALS	<i>Collaboration Recommender for Academic Social Networks</i>
FI	Filtragem de Informação
MAP	<i>Mean Average Precision</i>
SNA	<i>Social Network Analysis</i>
SR	Sistema de Recomendação
RI	Recuperação de Informação
RS	Rede Social
ST	Série Temporal
TSA	<i>Temporal Series Analysis</i>
XML	<i>Extensible Markup Language</i>

LISTA DE FIGURAS

Figura 2.1 Série temporal que mostra as vendas mensais (em quilolitros) de vinho tinto produzido por vinicultores australianos entre janeiro de 1980 e outubro de 1991.	28
Figura 2.2 Decomposição de uma série temporal em componentes de tendência, sazonalidade e ruído.....	29
Figura 4.1 Processo de recomendação utilizando a abordagem proposta, destacada pelo contorno tracejado.....	39
Figura 4.2 Exemplo 1: STs de Cooperação (esquerda) e Correlação (direita) entre os autores <i>A</i> e <i>B</i>	40
Figura 4.3 Exemplo 2: STs de Cooperação (esquerda) e Correlação (direita) entre os autores <i>C</i> e <i>D</i>	40
Figura 5.1 Processo de extração e organização dos dados dos pesquisadores e publicações. .	42
Figura 5.2 Gráfico mostrando a quantidade de publicações de cada autor (a) e por ano (b) da base de dados utilizada nos experimentos.	43
Figura 5.3 Gráfico mostrando a quantidade de publicações por área da base de dados utilizada nos experimentos.....	44

LISTA DE TABELAS

Tabela 2.1 Relação entre cooperação e correlação na decisão de reforçar uma colaboração existente.....	26
Tabela 5.1 Áreas de pesquisa e a quantidade de publicações para cada uma.	45
Tabela 5.2 Médias dos valores das métricas de avaliação para cada abordagem em relação à recomendação de colaborações que devem ser reforçadas.	48
Tabela 5.3 Médias dos valores das métricas de avaliação para cada abordagem em relação à recomendação de colaborações que devem ser iniciadas.....	49
Tabela 5.4 Resultados dos Testes-T executados para comparar a abordagem proposta com os outros algoritmos para recomendações de reforço de colaborações.	49
Tabela 5.5 Resultados dos Testes-T executados para comparar a abordagem proposta com os outros algoritmos para recomendações de iniciação de colaborações.....	50
Tabela 5.6 Valores das métricas Cobertura de Catálogo, Novidade e Diversidade para cada abordagem em relação à recomendação de colaborações que devem ser reforçadas.	50
Tabela 5.7 Valores das métricas Cobertura de Catálogo, Novidade e Diversidade para cada abordagem em relação à recomendação de colaborações que devem ser iniciadas. .	51

SUMÁRIO

1 INTRODUÇÃO	10
2 FUNDAMENTAÇÃO TEÓRICA	13
2.1 Análise de Redes Sociais	13
2.1.1 Conceitos.....	13
2.1.2 Aplicações.....	15
2.2 Sistemas de Recomendação	15
2.2.1 Abordagens	17
2.2.2 Avaliações de desempenho	18
2.3 Redes Sociais Acadêmicas e Recomendação de Colaborações	22
2.3.1 Conceitos.....	22
2.3.2 Métricas de recomendação.....	23
2.3.3 Funções de recomendação	25
2.4 Análise de Séries Temporais.....	27
2.4.1 Objetivos	28
2.4.2 Modelos.....	29
2.5 Resumo do Capítulo.....	30
3 TRABALHOS RELACIONADOS	32
3.1 Recomendações Sensíveis ao Tempo	32
3.2 Redes Acadêmicas.....	34
3.3 Resumo do Capítulo.....	36
4 RECOMENDAÇÃO DE COLABORAÇÕES ACADÊMICAS UTILIZANDO SÉ-	
RIES TEMPORAIS.....	37
4.1 Motivação e visão geral	37
4.2 Aplicação de séries temporais na previsão de métricas.....	38
4.3 Resumo do Capítulo.....	40
5 EXPERIMENTOS	41
5.1 Fonte de dados.....	41
5.1.1 Currículo Lattes	41
5.1.2 Definição das áreas das publicações	43
5.1.3 Casamento de coautorias das publicações	44
5.2 Geração de recomendações	46
5.3 Avaliações experimentais.....	47
5.3.1 Avaliação <i>offline</i>	48
5.3.2 Estudo de usuários	50
5.3.3 Análise dos resultados.....	53
5.4 Resumo do Capítulo.....	54
6 CONCLUSÃO	55
REFERÊNCIAS.....	57

1 INTRODUÇÃO

O avanço da tecnologia nos últimos anos permitiu a criação de Sistemas de Informação com acesso a grandes bases de dados, abrindo diversas possibilidades de aplicações. Tem-se como exemplo a Internet, onde uma enorme quantidade de dados é gerada e publicada a todo momento por usuários ao redor do mundo. Aos poucos foi surgindo e crescendo a necessidade de se filtrar essa massa de dados, de forma que as informações nela contida fossem úteis para os usuários. Foram então desenvolvidos métodos e técnicas de Recuperação de Informação (RI) e Filtragem de Informação (FI), que são estudados nas áreas de pesquisa em Mineração de Dados, em especial Mineração na Web (*Web Mining*) e, mais recentemente, em *Big Data* e *Web Science*.

Nesse contexto, foram desenvolvidos sistemas que ao longo do tempo passaram a ser conhecidos por Sistemas de Recomendação (SRs). Focados em resolver problemas de tomada de decisão, os SRs auxiliam o usuário a escolher um ou mais itens dentre vários em uma coleção de interesse. Por exemplo, são muito utilizados por empresas de comércio eletrônico a fim de personalizar a experiência de cada usuário ao acessar lojas online, garantindo que sejam oferecidos produtos do interesse do usuário (JANNACH et al., 2010). Porém, SRs também são encontrados em muitos outros contextos: desde recomendações de filmes, livros e músicas, até restaurantes, cinemas e outros lugares próximos geograficamente do usuário, entre outros (SCHAFER; KONSTAN; RIEDL, 1999). Exemplos atuais de SRs incluem Netflix¹, Amazon.com² e Spotify³.

A origem dos SRs também pode ser associada a trabalhos desenvolvidos nas áreas de ciência cognitiva, teoria de aproximação, teorias de previsão e ciências de gestão, porém eles emergiram como uma área de pesquisa independente na década de 90, quando os pesquisadores começaram a focar em problemas de recomendação com dependência nas estruturas de classificação (ADOMAVICIUS; TUZHILIN, 2005). Inicialmente, sites de comércio eletrônico empregaram SRs de forma a recomendar para seus clientes os produtos mais vendidos tendo como base uma classificação geral, os dados demográficos do cliente, ou uma análise das compras passadas do cliente (SCHAFER; KONSTAN; RIEDL, 1999). Porém, com o crescente interesse pela Análise de Redes Sociais (*Social Network Analysis* - SNA), surgiram métodos mais elaborados, utilizando a abordagem de filtragem de informação colaborativa, onde as recomendações feitas a um usuário levam em conta os produtos consumidos por outros usuários conectados de alguma

¹Netflix: <https://www.netflix.com/>

²Amazon: <https://www.amazon.com/>

³Spotify: <https://www.spotify.com/>

maneira a ele, seja diretamente, como um amigo ou familiar, ou indiretamente, como alguém que compartilha dos mesmos gostos no contexto em questão. Alguns exemplos de SRs que utilizam essa abordagem incluem LinkedIn⁴, Facebook⁵ e YouTube⁶.

Dentro desse contexto de SRs em redes sociais, alguns estudos vêm sendo realizados nos últimos anos a fim de explorar o potencial de SRs de coautoria de publicações e colaborações acadêmicas, visando identificar boas parcerias em diferentes áreas de estudo. É exatamente nesse contexto que este trabalho enquadra-se.

Alguns trabalhos já foram realizados com essa aplicação. Newman (2004) define uma abordagem de redes de coautoria em diferentes disciplinas, de maneira a permitir a análise de diversas questões sobre padrões de colaboração, tais como o número de artigos que um autor escreve, com quantas pessoas ele os escreve, qual é a distância típica entre os cientistas nestas redes e como os padrões de colaboração mudam entre diferentes assuntos e com o tempo. Lopes et al. (2010) utilizam um modelo de rede de coautoria similar para gerar recomendações de colaborações acadêmicas, através da análise das colaborações existentes e áreas em que os autores publicam artigos. Lopes (2012) explora mais profundamente esse contexto e propõe a primeira abordagem de recomendação de colaboração sensível ao tempo, utilizando pesos simples que valorizam as colaborações mais recentes, obtendo resultados satisfatórios. Outras abordagens de recomendações no contexto acadêmico são feitas em (BRANDÃO; MORO, 2012) e (BRANDÃO; MORO; ALMEIDA, 2013), e em outros contextos em (BASILE et al., 2015. Disponível em: <<http://ceur-ws.org/Vol-1388/DeCat2015-paper5.pdf>>. Acesso em: 5 abr. 2016), (CHEN et al., 2013), (GAO et al., 2013), (LEE; PARK; PARK, 2008) e (ULLAH et al., 2012), porém não foram encontrados outros trabalhos que explorassem o uso de informações temporais em redes acadêmicas.

O objetivo deste trabalho é incrementar a abordagem CORALS (LOPES, 2012) utilizando a Análise de Séries Temporais com a finalidade de melhorar as recomendações de colaborações no contexto acadêmico. Para isso, foram feitos dois experimentos: um experimento *offline* sem interação de usuários, e um estudo de usuário com a participação de um pequeno grupo de pesquisadores voluntários.

O restante da dissertação está organizado da seguinte forma:

- No Capítulo 2 são apresentados os conceitos básicos para um melhor entendimento deste trabalho. São apresentados conceitos sobre redes sociais, sistemas de recomendação,

⁴LinkedIn: <https://www.linkedin.com/>

⁵Facebook: <https://www.facebook.com/>

⁶YouTube: <https://www.youtube.com/>

recomendação de colaborações acadêmicas e séries temporais.

- No Capítulo 3 são apresentados alguns trabalhos relacionados e como eles se comparam com a abordagem proposta neste trabalho. São descritos brevemente alguns trabalhos sobre métodos de recomendação sensíveis ao tempo e recomendações no contexto acadêmico.
- No Capítulo 4 é apresentada a abordagem proposta para recomendação de colaborações no contexto acadêmico utilizando STs. Primeiro a motivação é apresentada e em seguida o método é descrito em detalhes. Por último, são descritos os experimentos realizados e são apresentadas as análises feitas sobre seus resultados de forma a comparar as abordagens existentes com a abordagem proposta.
- Por fim, no Capítulo 6 são apresentadas as considerações finais, destacando as contribuições e os resultados obtidos nos experimentos. Também são mencionadas as lições aprendidas no desenvolvimento deste trabalho e os caminhos possíveis para futuras investigações.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta um estudo bibliográfico a respeito das disciplinas abordadas nesta dissertação. São apresentados conceitos sobre Análise de Redes Sociais (Seção 2.1), Sistemas de Recomendação (Seção 2.2), Redes Sociais Acadêmicas e Recomendação de Colaborações (Seção 2.3), e Análise de Séries Temporais (Seção 2.4). Também são mencionados como cada área está relacionada com o presente trabalho e quais métodos e técnicas descritos são utilizados.

2.1 Análise de Redes Sociais

A Análise de Redes Sociais (*Social Network Analysis - SNA*) tem como base a importância das relações entre as suas unidades de interação. Os principais objetivos da análise de redes são medir e representar as relações estruturais de forma precisa, e explicar o porquê da ocorrência delas e quais são suas consequências (KNOKE; YANG, 2008). Wasserman e Faust (1994) definem SNA como uma perspectiva distinta nas Ciências Sociais e Comportamentais, envolvendo teorias, modelos e aplicações que são expressos em termos de conceitos e processos relacionais. Eles destacam os seguintes pontos:

- Atores e suas ações são vistos como interdependentes ao invés de unidades autônomas independentes;
- Laços relacionais entre atores são canais para a transferência de recursos materiais ou imateriais;
- Modelos de redes focados em indivíduos enxergam o ambiente estrutural da rede como fornecedor de oportunidades ou limitador de ação individual e;
- Modelos de redes conceituam estruturas (sociais, econômicas, políticas, etc.) como padrões duradouros de relacionamentos entre atores.

Wasserman e Faust (1994) também destacam o fato de que a unidade de análise da SNA não é o indivíduo, e sim uma entidade constituída de uma coleção de indivíduos e os laços entre eles, por isso é necessário empregar métodos especiais.

2.1.1 Conceitos

Há diversos conceitos-chave na SNA que são fundamentais para a compreensão e discussão de redes sociais. São eles: ator, vínculo relacional, díade, relação e rede social. A seguir

são apresentadas breves descrições desses conceitos, baseadas nas definições apresentadas em (WASSERMAN; FAUST, 1994).

Atores. Atores são as entidades cujos vínculos são estudados pela SNA. Atores podem ser indivíduos discretos ou unidades sociais coletivas. Exemplos de atores são pessoas em um grupo, departamentos de uma corporação, agentes de serviço público em uma cidade ou estado-nação no sistema mundial. Embora chamados de “atores”, essas unidades sociais não executam ações necessariamente. A maioria das aplicações de redes sociais lidam com coleções de atores que são todos do mesmo tipo, as chamadas redes de um modo (*one-mode networks*).

Vínculo relacional. Os vínculos relacionais são as conexões entre os atores da rede. Um vínculo estabelece uma conexão entre um par de atores e podem representar os mais variados conceitos. Alguns exemplos mais comuns são: avaliação de uma pessoa por outra (amizade, afeição, ou respeito), transferência de recursos materiais (transações comerciais ou emprestar coisas), associação ou afiliação (comparecer a um evento social juntos ou pertencer ao mesmo clube), interação comportamental (conversar e trocar mensagens), movimento entre lugares ou status (migração, mobilidade social ou física), conexão física (uma estrada, um rio ou uma ponte conectando dois pontos), relações formais (autoridade) e relação biológica (parentesco ou descendência).

Díade. Uma díade consiste em um par de atores e os possíveis laços entre eles. No nível mais básico, uma ligação ou relacionamento estabelece um laço entre dois atores. Esse laço é inerentemente uma propriedade do par e portanto não pode ser considerado algo próprio de um ator individualmente, por isso, os diversos tipos de SNA usam as díades como unidade de análise. As análises de díades focam nas propriedades das relações entre pares, como, por exemplo, se os laços são recíprocos ou não, ou se tipos específicos de múltiplas relações tendem a ocorrer juntas.

Relação. A coleção de laços de um tipo específico entre atores de um grupo é chamada de relação. Por exemplo, o conjunto de amizades entre crianças em uma classe, ou o conjunto de laços diplomáticos formais mantido entre pares de nações ao redor do mundo. Para qualquer grupo de atores podem ser medidas diversas relações diferentes, como, por exemplo, além dos laços diplomáticos formais entre as nações, pode-se ter a quantidade de dólares trocados entre elas durante um determinado ano.

Rede social. Considerando os conceitos definidos anteriormente, uma rede social é constituída por um conjunto ou conjuntos finitos de atores e pela relação ou relações definidas entre eles.

Triáde, subgrupo e grupo são outros conceitos na Análise de Redes Sociais definidas por

Wasserman e Faust (1994), porém não são utilizados neste trabalho e por isso suas definições foram omitidas.

2.1.2 Aplicações

Atualmente existem diversos tipos de redes sociais e, em geral, é possível utilizar a análise de redes sociais para se obter informações que não seriam possíveis de se conseguir com uma simples análise estatística. Algumas redes muito populares, como o Facebook, o LinkedIn e o ResearchGate¹, possuem mecanismos que identificam padrões nas conexões de seus usuários e os utiliza para melhorar a experiência deles, seja recomendando novas conexões ou selecionando informações diferenciadas para exibir ao usuário (como propagandas de produtos que provavelmente vão lhe interessar).

2.2 Sistemas de Recomendação

Sistemas de Recomendação (SRs) envolvem ferramentas e técnicas utilizadas para auxiliar usuários em processos de tomada de decisão, tais como quais produtos comprar, quais músicas escutar ou quais notícias ler (RICCI; ROKACH; SHAPIRA, 2011). Uma grande demanda por SRs surgiu nos últimos anos com o avanço da tecnologia, que permitiu a criação de Sistemas de Informação com acesso a grandes bases de dados. Tem-se como exemplo a Internet, onde uma enorme quantidade de dados é gerada e publicada a todo momento por usuários ao redor do mundo. Como consequência disso, surgiu a necessidade de desenvolver métodos e técnicas para extrair informações úteis desses repositórios. Esses métodos e técnicas são estudados na área de pesquisa em Mineração de Dados, em especial na Mineração na Web (*Web mining*), e, mais recentemente, em *Big Data* e *Web Science*.

Na década de 60 surgiram as primeiras abordagens de Recuperação da Informação (RI), a qual consiste na tarefa de “encontrar material de natureza não estruturada que satisfaça uma necessidade de informação dentro de grandes coleções” (MANNING et al., 2008). O foco principal da RI é a construção de índices que permitam consultas em grandes bases de dados. Por exemplo, em uma coleção de livros pode-se criar índices por assunto, título ou autor, para que um usuário possa consultar todos os livros escritos por Robin Cook ou todos os livros sobre Linguagens de Programação, e se possa encontrar os livros desejados rapidamente através desses índices. Essas bases são relativamente estáticas, ou seja, dados são adicionados ou removidos

¹<https://www.researchgate.net/>

da base com pouca frequência, enquanto as consultas são dinâmicas, ou seja, cada consulta efetuada é normalmente diferente da outra.

Nas décadas seguintes surgiu a Filtragem de Informações (FI), cujo objetivo é expor o usuário apenas a informações que sejam relevantes a ele (HANANI; SHAPIRA; SHOVAL, 2001). Diferente da RI, a FI consiste em buscar informações em bases de dados bastante dinâmicas, com base em uma consulta relativamente estática. Essas bases de dados são, na verdade, fluxos de dados recebidos por um usuário, e a consulta é formada por uma descrição dos seus interesses, e pode ser vista como um “perfil”. Por exemplo, ao acessar um site de notícias, um usuário pode se interessar apenas pelas notícias sobre tecnologia, e a quantidade de notícias indesejadas, sobre outros assuntos, pode ser muito grande, dificultando a navegação do usuário e fazendo-o perder muito tempo buscando o que lhe interessa. Assim, a FI evita que o usuário seja surpreendido pela explosão de conteúdo disponível, omitindo todas as informações indesejadas (notícias sobre política, educação, economia, etc.) para que o usuário possa focar no que lhe interessa (no caso, as notícias sobre tecnologia). De forma inversa, pode-se filtrar itens indesejados como, por exemplo, SPAM recebido via e-mail.

Os SRs são uma subclasse de sistemas de FI que busca prever a nota ou preferência que um usuário daria para um determinado item. Os elementos básicos de um SR são: usuários, itens e avaliações.

Usuários. São as pessoas que usam o sistema e possuem preferências por itens. Os usuários podem possuir algumas propriedades como, por exemplo, informações demográficas (se o usuário é homem ou mulher, possui menos ou mais de 18 anos, etc.).

Itens. São os elementos ou objetos que o sistema recomenda, como, por exemplo, músicas, notícias, livros, etc. Eles também podem possuir atributos que os caracterizem, como o gênero de um filme, o autor de um livro, etc. Em alguns contextos, os itens de recomendação podem ser os próprios usuários. Por exemplo, em uma rede social acadêmica, um SR pode recomendar pesquisadores (usuários) com interesses similares aos do usuário que o está utilizando, como é o caso da abordagem desenvolvida e avaliada neste trabalho.

Avaliações. São valores que conectam os usuários aos itens, através de uma representação de preferência de um usuário por um item. Essa representação pode ser uma nota entre 1 e 5 ou uma simples indicação se o usuário gostou ou não de um item. A partir dessas informações é possível construir um modelo de usuário, ou seja, uma descrição estruturada que considera as informações que foram coletadas de maneira explícita ou implícita a respeito do usuário durante sua interação com o sistema, incluindo as avaliações feitas por ele a respeito de itens.

SRs podem utilizar esse modelo de diversas formas. O YouTube possui itens que são vídeos

dos mais variados tipos, desde clipes musicais e *trailers* de filmes, até vídeo-aulas e *vlogs*. Para cada usuário cadastrado no sistema, ele armazena informações sobre os vídeos assistidos, os vídeos marcados como favoritos e quais vídeos o usuário explicitamente indicou ter gostado ou não. A partir desse perfil de usuário, o sistema oferece recomendações de novos vídeos que possam lhe interessar ou vídeos já assistidos que o usuário possa querer assistir novamente. O caso do Netflix é muito parecido, pois se trata de um serviço de *streaming* de filmes e seriados. O Netflix apresenta recomendações de novos filmes e seriados que o usuário ainda não assistiu a partir de informações de seu histórico, indicando qual filme ou seriado assistido recentemente foi decisivo para a recomendação ocorrer. Essa informação adicional pode ser crucial para o usuário aceitar ou não a recomendação, pois ele poderá ter uma boa noção do conteúdo que está sendo recomendado apenas recordando-se de algo que assistiu recentemente, tornando o processo de decisão mais fácil. O Facebook utiliza dados de preferências explícitas dos usuários para recomendar produtos e outros serviços através de um sistema inteligente de publicidade. Informações implícitas também são utilizadas, como, por exemplo, se um grupo de usuários conectados demonstrou algum interesse específico, usuários conectados a esse grupo podem compartilhar deste interesse.

2.2.1 Abordagens

Os SRs são classificados em três tipos: baseados em conteúdo, colaborativos e híbridos (ADOMAVICIUS; TUZHILIN, 2005). Essas abordagens são descritas a seguir.

Filtragem baseada em conteúdo. Analisa as descrições dos itens para identificar aqueles que podem ser de interesse para o usuário (PAZZANI; BILLSUS, 2007). Embora os detalhes de cada sistema variem, SRs baseados em conteúdo têm em comum o fato de adotarem alguma forma de descrição do item a ser recomendado, uma forma de criar o perfil dos usuários (descrição dos itens que eles gostam) e uma maneira de comparar itens que possam ser recomendados com os itens presentes no perfil do usuário, determinando o que recomendar. O perfil do usuário é frequentemente criado e atualizado automaticamente em resposta ao *feedback* em relação a quanto o usuário deseja (i.e., interessa-se) dos itens que foram apresentados a ele. Segundo Lops, Gemmis e Semeraro (2011), esta abordagem tem como vantagens a independência entre usuários (as recomendações são baseadas apenas no perfil do próprio usuário), a transparência (o sistema pode fornecer as características ou descrições dos itens que causaram a recomendação do mesmo) e a capacidade de recomendar itens novos que ainda não foram avaliados por nenhum usuário. Por outro lado, a forma de descrever os itens é limitada e pode não ser o su-

ficiente para permitir uma boa interpretação das preferências do usuário, além do problema de superespecialização que torna as recomendações pouco inovadoras, pois são limitadas aos itens similares aos que o usuário avaliou. Outro problema comum nesta abordagem é o *cold start*, quando um usuário novo no sistema ainda não forneceu *feedback* o suficiente para o sistema gerar recomendações. Esse problema costuma ser resolvido com abordagens híbridas que utilizam mais informações além das avaliações feitas pelo usuário, como informações demográficas do usuário e informações de outros usuários.

Filtragem colaborativa. É o processo de filtrar ou avaliar itens através da opinião de outras pessoas (SCHAFER et al., 2007). Tanto os usuários quanto os itens possuem conjuntos de avaliações. Os usuários possuem avaliações em relação aos itens e os itens possuem avaliações feitas por usuários a respeito deles. Quando combinados, esses conjuntos formam uma matriz esparsa de avaliações, onde a tarefa de prever valores é equivalente a preencher as células em branco, enquanto que a tarefa de recomendar consiste em selecionar células promissoras, ou seja, que foram preenchidas com boas avaliações pelo algoritmo. Lika, Kolomvatsos e Hadjiefthymiadis (2014) propõem uma abordagem que utiliza os dados demográficos de novos usuários para conectá-los a usuários existentes que já avaliaram itens anteriormente, de forma a resolver o problema de *cold start* com novos usuários.

Filtragem Híbrida. Um SR híbrido é aquele que combina a filtragem baseada em conteúdo e a colaborativa, afim de alcançar uma sinergia entre elas (BURKE, 2007). Esta abordagem permite tirar vantagem dos pontos fortes de cada técnica e minimizar os pontos fracos das mesmas, dependendo da forma que for construída.

Há ainda outras classificações de SRs que levam em consideração a fonte dos dados utilizados pelo sistema, como a abordagem baseada em conhecimento e a demográfica, sendo detalhadas em (BURKE, 2002) e (BOBADILLA et al., 2013).

O presente trabalho utiliza uma abordagem híbrida que combina o uso de atributos dos pesquisadores para calcular a correlação entre eles (filtragem baseada em conteúdo) e a análise dos seus relacionamentos para calcular a cooperação entre eles (filtragem colaborativa). Esta abordagem é definida por Lopes (2012) e é utilizada como base para a abordagem sensível ao tempo que este trabalho propõe.

2.2.2 Avaliações de desempenho

Shani e Gunawardana (2011) definem três tipos de experimentos para avaliação de SRs: os experimentos *offline*, os estudos de usuários e os experimentos *online*.

Experimentos *Offline*. São feitos com conjuntos de dados coletados ou gerados previamente. Usa dados que descrevem escolhas ou avaliações de usuários a respeito de uma seleção de itens para simular o comportamento deles ao interagir com o SR. A principal vantagem deste tipo de experimento é que não é necessária interação com usuários reais, permitindo que se comparem diversos algoritmos diferentes a um baixo custo. Por outro lado, experimentos *offline* permitem que se avalie apenas uma parcela das propriedades de um sistema, como, por exemplo, o poder de previsão dos algoritmos, sem considerar a influência do sistema no comportamento do usuário. Por esses motivos, considera-se esta uma abordagem de avaliação inicial para selecionar os melhores algoritmos a serem testados nos estudos de usuários e experimentos *online*.

Estudos de Usuário. São experimentos conduzidos com um conjunto de usuários que são orientados a executar diversas tarefas envolvendo interação com o SR. Enquanto os usuários interagem com o sistema, pode-se observar e registrar o comportamento deles, como quais tarefas foram executadas, a exatidão dos resultados das tarefas ou o tempo que cada usuário levou para completá-las. Em vários casos também se pode pedir que os usuários respondam questionários qualitativos antes, durante ou após a tarefa ser executada. Esses questionários podem coletar dados que não são diretamente observáveis, tais como se o usuário achou a interface do sistema agradável ou se achou a tarefa fácil de ser completada. Estudos de usuário podem ser utilizados quando experimentos *offline* não são aplicáveis, mas também podem complementá-los. Porém, são mais caros de se conduzir, pois dependem da disponibilidade de usuários e das tarefas a serem executadas, impedindo que se possa testar todos os cenários. Estudos de usuário são a abordagem que permite responder a uma quantidade mais ampla de questões, pois, diferente de experimentos *offline*, permite que se teste o comportamento do usuário ao interagir com o sistema, além de ser a única abordagem que permite a coleta de dados qualitativos, que são muitas vezes essenciais para a interpretação de dados quantitativos (LOPS; GEMMIS; SEMERARO, 2011).

Experimentos *Online*. Envolvem o uso do SR por usuários reais, que estão geralmente alheios ao experimento, executando tarefas reais. Embora exista um limite dos dados que podem ser extraídos deste tipo de experimento, é o que mais se aproxima da realidade. Uma desvantagem é que ao executar experimentos *online* corre-se o risco de fazer recomendações fracas, o que levaria os usuários a deixarem de usar o sistema. Por isso, é recomendado que se execute estes experimentos por último, depois dos experimentos *offline* que forneçam evidências de que as abordagens utilizadas são razoáveis e, talvez, após estudos de usuário para medir a atitude dos usuários frente ao sistema.

Neste trabalho foram executados o experimento *offline* e o estudo de usuários. O experimento *offline* permitiu uma análise preliminar da comparação entre os algoritmos que serviram de base para este trabalho e o algoritmo proposto, e permite fácil reprodução e refinamento em períodos de tempo curtos. Já o estudo de usuários permite uma análise preliminar do comportamento dos usuários frente às recomendações feitas. Por último deve ser feito um experimento *online* com um maior número de usuários e um período de tempo mais longo, porém não foi possível realizar este experimento por limitações de tempo.

Após os experimentos é necessário processar os dados coletados a fim de avaliar os métodos utilizados para gerar recomendações. Para isso, usa-se uma série de métricas provenientes da RI, das quais Precisão (*Precision*), Revocação (*Recall*), Média das Precisões Médias (*Mean Average Precision - MAP*) e Precisão até 10 (*Precision at 10 - Pr@10*) são bastante utilizadas no contexto de SRs e foram escolhidas para avaliar a abordagem proposta deste trabalho. As métricas são descritas a seguir.

Precisão. É a fração das recomendações feitas que são relevantes para o usuário. É calculada de acordo com a Equação 2.1.

$$\text{Precisão} = \frac{|\text{itens relevantes} \cap \text{itens recuperados}|}{|\text{itens recuperados}|} \quad (2.1)$$

Média das Precisões Médias. É a média das precisões médias de um conjunto de usuários. Precisão média é a média das precisões computadas para cada ponto de recomendações relevantes na sequência do *ranking*. Esta métrica leva em conta a posição das recomendações relevantes, valorizando as mais bem posicionadas.

Precisão até K . A Precisão até K para um conjunto de usuários é a média das precisões calculadas até a posição K do *ranking* de resultados. Um dos valores mais usados para K é 10.

Revocação. É a fração das recomendações relevantes para o usuário que foram recuperadas com sucesso. É calculada através da Equação 2.2.

$$\text{Revocação} = \frac{|\text{itens relevantes} \cap \text{itens recuperados}|}{|\text{itens relevantes}|} \quad (2.2)$$

Além dessas métricas, as diferentes abordagens são comparadas utilizando o **Teste-T**, que é um teste estatístico empregado para verificar a existência de diferenças estatisticamente significativas. Foi utilizado o Teste-T bicaudal com o p-valor 5

McNee, Riedl e Konstan (2006) defendem que as recomendações mais precisas talvez não sejam as mais úteis para os usuários, e que as métricas de avaliação deveriam levar em conta outros fatores que impactam na qualidade das recomendações (como “Serendipidade” - ver

a seguir), e serem aplicadas em listas de recomendações ao invés de itens individuais. Ge, Delgado-Battenfeld e Jannach (2010) defendem o uso das métricas Cobertura (*Coverage*) e “Serendipidade” (*Serendipity*).

Cobertura. A Cobertura de um SR é a medida dos itens do domínio do sistema em que o mesmo pode fazer previsões ou recomendações (HERLOCKER et al., 2004). O termo é associado a dois conceitos: a percentagem de itens para os quais o sistema é capaz de gerar recomendações (cobertura de previsão) e a percentagem de itens disponíveis que são efetivamente recomendados a um usuário (cobertura de catálogo). Ge, Delgado-Battenfeld e Jannach (2010) definem as Equações 2.3 e 2.4 para calcular as coberturas de previsão e catálogo, respectivamente. No caso, I é o conjunto de todos os itens do sistema; I_p é o conjunto de todos os itens os quais o sistema pode gerar recomendações; e I_L^j é o conjunto de itens contidos na lista L retornada na j -ésima recomendação observada durante as medições.

$$\text{Cobertura de previsão} = \frac{|I_p|}{|I|} \quad (2.3)$$

$$\text{Cobertura de catálogo} = \frac{|\bigcup_{i=1}^N I_L^i|}{|I|} \quad (2.4)$$

Nas avaliações deste trabalho utilizou-se apenas a Cobertura de Catálogo, porque considera-se que os algoritmos são capazes de gerar recomendações para todos os itens, já que eles foram selecionados por terem atributos (publicações) o suficiente para incorporar os experimentos.

Serendipidade. A Serendipidade é a medida da extensão em que os itens recomendados são atraentes e surpreendentes ao usuário (HERLOCKER et al., 2004). De acordo com Ge, Delgado-Battenfeld e Jannach (2010), um item com alta serendipidade deve ser um item que o usuário ainda não descobriu e não espera que exista, mas que deve ser interessante, relevante e útil para o ele. Ainda não há uma forma completamente aceita para o cálculo desta métrica, porém, Brandão e Moro (2012) propõem uma métrica adaptada para o contexto de redes de colaboração acadêmica chamada *Diversidade*, descrita a seguir.

Brandão e Moro (2012) introduzem duas métricas de avaliação adaptadas para o contexto das redes de colaboração acadêmica: Novidade e Diversidade, descritas a seguir.

Novidade. São consideradas “novas recomendações” as indicações de itens que o usuário ainda não conhece ou não teria conhecimento sem o uso do SR. Brandão e Moro (2012) adaptaram a proposta de Fouss e Saerens (2008), que mede o “novo” e o “original” presentes em uma lista de recomendações, para o contexto de redes de colaboração acadêmica. Primeiro são definidas as frequências dos itens recomendados entre todos os usuários e é calculada a mediana

destas frequências. Depois essa mediana é dividida pelo número total de usuários, resultando em um valor entre 0 e 1. Um resultado alto significa que o algoritmo recomenda itens populares, que são bem conhecidos e portanto não resultam em novas recomendações, por isso um valor baixo indica uma Novidade melhor. Portanto, o valor final de Novidade é definido como o complemento deste valor.

Diversidade. Uma lista de recomendações é dita diversa quando a similaridade entre os seus itens é baixa. Um dos métodos mais utilizados para isso é a similaridade entre cada par de itens da lista (SHANI; GUNAWARDANA, 2011). Em geral, essa similaridade é definida pela correlação de Pearson ou distância de cosseno (ZIEGLER et al., 2005). No contexto de redes de colaboração acadêmica, Brandão e Moro (2012) utilizam os valores de Correlação para determinar a similaridade entre os pesquisadores, já que essa métrica representa a similaridade entre dois pesquisadores de acordo com as áreas em que atuam. O valor final da Diversidade de uma lista de recomendações é o complemento da média das similaridades de cada par de itens da lista. Define-se como Diversidade de um algoritmo a média das Diversidades das listas de recomendação para cada usuário.

Além das métricas provenientes da RI, as métricas de Cobertura (de Catálogo), Novidade e Diversidade são utilizadas neste trabalho durante a análise comparativa dos algoritmos no experimento *offline* (Seção 5.3.1).

2.3 Redes Sociais Acadêmicas e Recomendação de Colaborações

Nesta seção são descritos os principais conceitos de Redes Sociais Acadêmicas (Subseção 2.3.1), seguidos das métricas utilizadas nas funções de recomendação de colaboração (Subseção 2.3.2) e as próprias funções (Subseção 2.3.3).

2.3.1 Conceitos

O contexto deste trabalho considera uma rede de colaborações acadêmicas, formada por pesquisadores que atuam em diversas áreas. Para a melhor compreensão dessa rede, são definidos os conceitos de “pesquisadores”, “publicações” e “áreas de pesquisa”.

Pesquisadores. Pesquisadores são os atores de uma rede social acadêmica e os usuários de um SR de colaborações. Eles são conectados diretamente caso tenham publicado algum trabalho juntos (NEWMAN, 2004).

Publicações. As publicações podem ser vistas como atributos dos laços existentes entre

pesquisadores. Dois pesquisadores podem ter publicado diversos trabalhos juntos, porém são ligados por apenas um laço. Publicações com mais de dois autores podem aparecer como atributos de laços que conectam diferentes pares de pesquisadores.

Áreas de pesquisa. Cada publicação possui a informação de área de pesquisa com a qual está relacionada. Esta informação serve para agrupar os trabalhos com foco comum. A partir dessa associação, é possível determinar quais pesquisadores trabalham nas mesmas áreas.

2.3.2 Métricas de recomendação

A abordagem CORALS, proposta por Lopes (2012), define três métricas básicas para gerar recomendações no contexto das redes de colaboração: “Cooperação”, “Proximidade Social” e “Correlação”. Além dessas métricas, também é definido um peso sensível ao tempo e o mesmo é aplicado à métrica de Cooperação de duas formas diferentes, chamadas de TR e TG, descritas em seguida.

Cooperação. A Cooperação ($Cp_{i,j}$) representa o grau de cooperação entre dois autores i e j . Esta métrica estima um valor decimal entre 0 e 1, e representa a proporção de trabalhos produzidos pela colaboração de dois autores i e j ($p_{i,j}$) sobre o total de trabalhos produzidos pelo autor i (p_i), como descrito pela equação 2.5.

$$Cp_{i,j} = \frac{p_{i,j}}{p_i} \quad (2.5)$$

Proximidade Social. A proximidade social é definida por uma variante normalizada do método do menor caminho. No presente contexto, a métrica de cooperação foi utilizada como base para o cálculo das distâncias entre os nodos (pesquisadores). A Equação 2.6 determina o peso de uma ligação entre dois pesquisadores.

$$d_{i,j} = \begin{cases} \infty, & \text{se } (Cp_{i,j} = 0) \\ 1 + (1 - Cp_{i,j}), & \text{caso contrário} \end{cases} \quad (2.6)$$

A Equação 2.7 determina a proximidade propriamente dita, que é assimétrica, ou seja, um pesquisador i pode ter um valor maior de proximidade com um pesquisador j do que o pesquisador j tem com i .

$$Sc_{i,j} = \begin{cases} 0, & \text{se } (shortest_path(i, j) = \infty) \\ \frac{(MAX+1)-shortest_path(i,j)}{MAX}, & \text{caso contrário} \end{cases} \quad (2.7)$$

$shortest_path(i, j)$ corresponde ao menor caminho entre os pesquisadores i e j na direção $i \rightarrow j$, utilizando a distância d como peso das relações. MAX é uma função que retorna o valor do máximo menor caminho encontrado entre dois pesquisadores da rede em questão, ou seja, $MAX = shortest_path(x, y)$, onde x e y formam o par de pesquisadores cujo menor caminho entre eles, na direção $x \rightarrow y$, é o maior de todos os menores caminhos entre todos os pares de pesquisadores, em todas as direções.

Correlação. A Correlação considera as diferentes áreas de pesquisa nas quais os pesquisadores atuam. É construído um perfil de atuação para cada pesquisador, onde são associados pesos para representar o grau de atuação dos mesmos em diferentes áreas de pesquisa, em relação a seus perfis globais. A Equação 2.8 calcula o peso de cada área que irá compor o perfil de atuação do pesquisador. O peso $R_{i,k}$ de um pesquisador i em uma área de pesquisa k , é determinado pela divisão de $p_{i,k}$, o número total de trabalhos publicados pelo pesquisador i na área k , por p_i , o total de trabalhos publicados pelo pesquisador i .

$$R_{i,k} = \frac{p_{i,k}}{p_i} \quad (2.8)$$

Após a construção dos perfis de atuação de cada pesquisador, a correlação entre eles é calculada de acordo com a Equação 2.9.

$$Cr_{i,j} = \frac{\sum_{k=1}^n R_{i,k} \cdot R_{j,k}}{\sqrt{\sum_{k=1}^n (R_{i,k})^2 \cdot \sum_{k=1}^n (R_{j,k})^2}} \quad (2.9)$$

Pesos com fator temporal

Lopes (2012) apresenta a única abordagem conhecida que considera a influência de aspectos temporais para determinar os pesos relacionais em uma rede social que é utilizada para recomendação de colaborações acadêmicas. Os aspectos temporais considerados estão relacionados ao ano das publicações, que é usado para inferir o quão recentes são as colaborações. Dessa forma, colaborações mais recentes são representadas por valores maiores.

São definidos dois pesos sensíveis ao tempo: TR e TG . A diferença entre os dois pesos é descrita a seguir. Ambos utilizam como base o fator temporal descrito pela Equação 2.10.

$$t_k = \begin{cases} \frac{w - (y_r - y_k)}{w}, & \text{se } (y_r - y_k) < w \\ t_{min}, & \text{caso contrário} \end{cases} \quad (2.10)$$

w é o intervalo de tempo a ser depreciado proporcionalmente (do ano de publicação mais recente para o mais antigo), y_r é o ano de publicação mais recente presente na rede de colaborações, y_k é o ano de publicação para o qual o fator temporal está é calculado, e t_{min} é o valor mínimo de t_k que pode ser gerado e aplicado para os anos cuja diferença é maior ou igual a w .

A partir desse fator temporal, define-se a métrica **TR**, descrita pela Equação 2.11. Essa métrica aplica o fator temporal para o ano de publicação mais recente entre dois pesquisadores. Dessa forma, os relacionamentos que ocorreram mais recentemente são mais valorizados. A métrica é definida pela multiplicação simples da métrica de cooperação com o fator temporal.

$$Tr_{i,j} = Cp_{i,j} \cdot t_{y_{i,j}} \quad (2.11)$$

Outra métrica definida a partir do fator temporal é a **TG**. Diferente da métrica TR, esta métrica aplica o fator temporal de forma global, ou seja, para cada ano em que houve publicações dos pesquisadores em questão.

$$Tg_{i,j} = \frac{\sum_{k=1}^{n_{i,j}} t_{y_k}}{n_i \sum_{x=1} t_{y_x}} \quad (2.12)$$

A Equação 2.12 é definida pela redefinição da métrica de cooperação. Porém, ao invés de considerar a quantidade de trabalhos em colaboração entre dois pesquisadores pelos trabalhos totais de um pesquisador, esses valores são substituídos pelo somatório dos fatores temporais calculados para cada uma das publicações consideradas.

A abordagem sensível ao tempo proposta por este trabalho utiliza as métricas de Cooperação, Proximidade Social e Correlação como base, fazendo previsões sobre séries temporais de observações destas métricas ao longo do tempo. As métricas TR e TG não são utilizadas em conjunto com as séries temporais, mas foram incluídas na análise comparativa feita no experimento *offline*.

2.3.3 Funções de recomendação

Após o cálculo das métricas, são definidas duas possíveis ações para cada recomendação encontrada: iniciar colaboração quando dois autores não publicaram juntos anteriormente ou intensificar a colaboração quando eles já publicaram juntos.

A Equação 2.13 define um score como uma média ponderada entre as métricas de Cor-

relação e Proximidade Social, e é utilizada a seguir como escore para definir o *ranking* das recomendações para iniciar colaborações. Os pesos w_{Cr} e w_{Sc} podem ser usados para enfatizar uma das métricas utilizadas, e podem ser definidas através de experimentação.

$$Cr_{Sc_{i,j}} = \frac{w_{Cr} \cdot Cr_{i,j} + w_{Sc} \cdot Sc_{i,j}}{w_{Cr} + w_{Sc}} \quad (2.13)$$

Para o caso das colaborações existentes que devem ser reforçadas, as relações de Cooperação e Correlação são combinadas como mostra a Equação 2.14.

$$Cp_{Cr_{i,j}} = \frac{Cp_{i,j}}{Cr_{i,j}} \quad (2.14)$$

A Equação 2.15 descreve o processo utilizado para definir qual ação irá se tomar. Para a recomendação de iniciação de colaboração ocorrer é necessário que a cooperação entre os pesquisadores não exista e o escore $Cr_{Sc_{i,j}}$ seja maior que um limiar previamente definido. Para a recomendação de reforço de colaboração ocorrer, as métricas de cooperação e correlação devem seguir um padrão pré-definido, descrito pela Tabela 2.1. Considera-se que as métricas são “baixas” (*low*) quando os valores medidos forem inferiores a 1/3, “médias” (*medium*) quando estiverem entre 1/3 e 2/3, e “altas” (*high*) quando forem superiores a 2/3. Caso não se atinjam os requisitos para uma das duas ações, a recomendação não é feita.

$$Action_{i,j} = \begin{cases} Initiate_Collaboration, & \text{se } (Cp_{i,j} = 0) \wedge \\ & (Cr_{Sc_{i,j}} > threshold) \\ Intensify_Collaboration, & \text{se } (Cp_{i,j} \in low) \wedge \\ & ((Cr_{i,j} \in medium) \vee \\ & (Cr_{i,j} \in high)) \\ Not_Recommend, & \text{caso contrário} \end{cases} \quad (2.15)$$

Tabela 2.1 – Relação entre cooperação e correlação na decisão de reforçar uma colaboração existente.

		Correlação		
		Baixa	Média	Alta
Cooperação	Baixa	OK	Recomendar	Recomendar
	Média	Alerta	OK	Alerta
	Alta	Alerta	Alerta	OK

Fonte: retirada de (LOPES, 2012)

Determinada qual ação irá ser tomada, são utilizados os valores calculados pelas métricas $Cr_{Sc_{i,j}}$ e $Cp_{Cr_{i,j}}$ para determinar os escores de cada recomendação, de acordo com a Equação 2.16. Elas também são separadas em duas listas, de acordo com a ação a ser tomada para cada uma, e ordenadas. Para o caso de recomendações de iniciação, elas são ordenadas de forma decrescente, uma vez que quanto maior o valor da métrica, maior a possibilidade de a colaboração ser interessante para o usuário-alvo da recomendação. Para o caso de recomendações de reforço de uma colaboração existente, a lista é ordenada de forma crescente, uma vez que as recomendações são apresentadas em ordem de necessidade de intensificação.

$$Score_{i,j} = \begin{cases} Cr_{Sc_{i,j}}, & \text{se } (Action_{i,j} = \\ & \quad Initiate_Collaboration) \\ Cp_{Cr_{i,j}}, & \text{se } (Action_{i,j} = \\ & \quad Intensify_Collaboration) \\ 0, & \text{caso contrário} \end{cases} \quad (2.16)$$

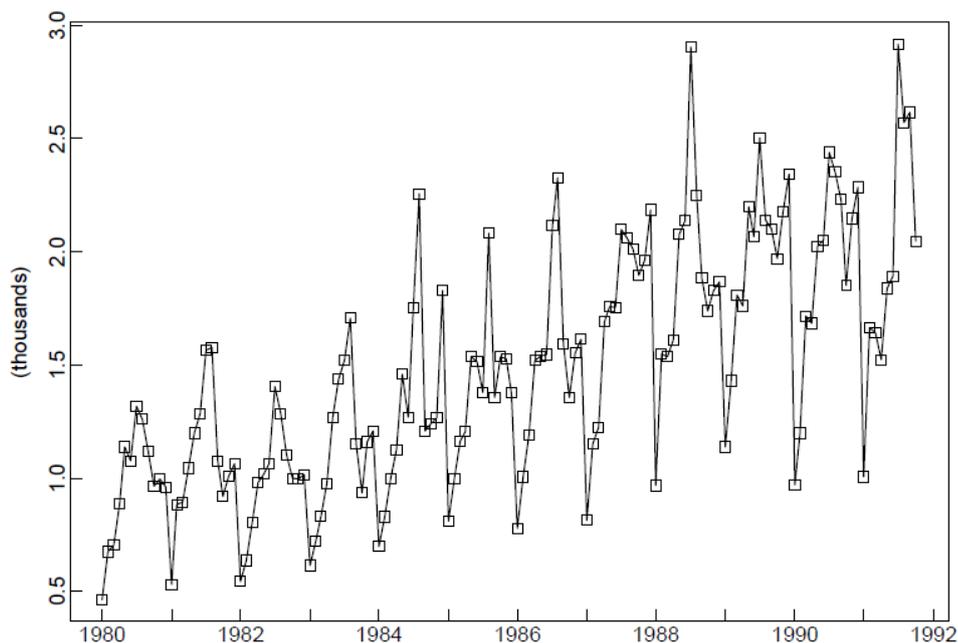
2.4 Análise de Séries Temporais

De acordo com Brockwell e Davis (2006), uma Série Temporal (ST) é um conjunto de observações feitas ao longo do tempo. Uma ST pode ser discreta, se as observações forem feitas em um conjunto discreto de pontos no tempo, ou contínua, se as observações forem feitas em um período contínuo de tempo. A Figura 2.1 mostra um exemplo de ST com dados de vendas mensais de vinho tinto ao longo do tempo.

Componentes. As STs podem ser decompostas em até 4 componentes diferentes:

- **Tendência.** Pode ser definida como mudanças de longo prazo no nível da média. Por exemplo, variáveis climáticas algumas vezes exibem variações cíclicas no decorrer de períodos longos de tempo, tal como 50 anos.
- **Ciclo.** Longas ondas, mais ou menos regulares, em torno de uma linha de tendência. Por exemplo, as variações diárias de temperatura.
- **Sazonalidade.** Representa padrões regulares de alteração de curto-prazo no nível da média. Por exemplo, séries com variações anuais, como a quantidade de vendas ao longo do ano ou leituras de temperatura.
- **Ruído.** Resíduo não incorporado pelos outros componentes que descrevem influências irregulares e aleatórias.

Figura 2.1 – Série temporal que mostra as vendas mensais (em quilolitros) de vinho tinto produzido por vinicultores australianos entre janeiro de 1980 e outubro de 1991.



Fonte: retirada de (BROCKWELL; DAVIS, 2006)

A Figura 2.2 mostra uma ST com dados de nascimentos por mês na cidade de Nova York ao longo dos anos e seus componentes isoladas de tendência, sazonalidade e ruído.

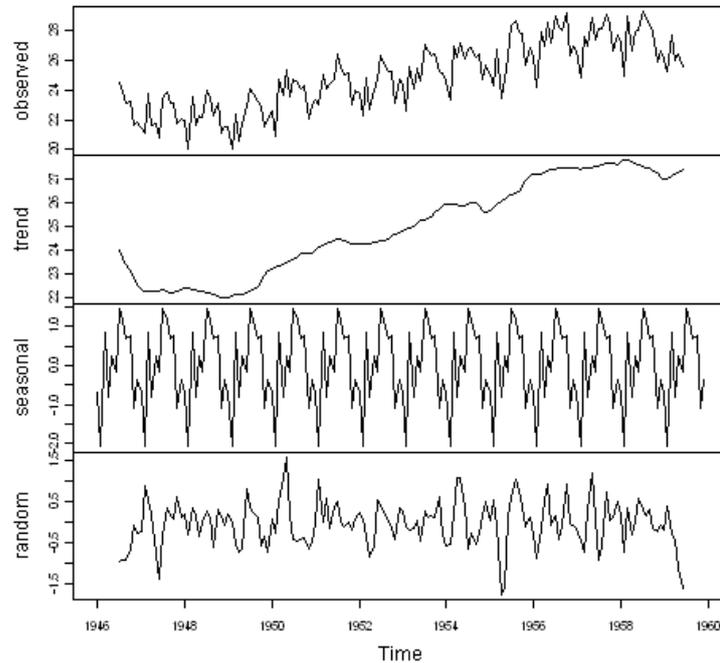
2.4.1 Objetivos

A Análise de Séries Temporais (*Temporal Series Analysis - TSA*) é composta por métodos para analisar dados de STs de forma a extrair estatísticas significativas e outras características dos dados. Os objetivos da TSA podem ser divididos em quatro grupos: descrever, explicar, prever e controlar (CHATFIELD, 2013).

Descrever. O primeiro passo da análise geralmente consiste em gerar um gráfico com as observações para se obter o chamado *plot* de tempo (*time plot*), e então obter métricas descritivas simples das principais propriedades da série. Por exemplo, a Figura 2.1 mostra o número de vendas de vinho tinto ao longo dos anos, onde se pode perceber padrões como o aumento das vendas durante o inverno e a redução das mesmas no verão.

Explicar. Quando observações são feitas sobre duas ou mais variáveis, pode ser possível utilizar a variação em uma ST para explicar a variação de outra ST. Isso pode levar a um entendimento mais profundo do mecanismo que gera determinada ST. Por exemplo, é interessante ver como o nível do mar é afetado pela temperatura e pressão, e ver como vendas são afetadas

Figura 2.2 – Decomposição de uma série temporal em componentes de tendência, sazonalidade e ruído.



Fonte: retirada de Coghlan (2014. Disponível em: <<https://media.readthedocs.org/pdf/a-little-book-of-r-for-time-series/latest/a-little-book-of-r-for-time-series.pdf>>. Acesso em: 5 apr. 2016.)

pelo preço e condições econômicas.

Prever. Dada uma ST, pode ser desejado prever os valores futuros da série. Por exemplo, prever como as vendas irão se comportar nos próximos meses, ou como a economia irá se comportar. Esse é o objetivo que será explorado neste trabalho, descrito no Capítulo 4.

Controlar. STs também são utilizadas para melhorar o controle sobre algum sistema físico ou econômico. Por exemplo, uma ST pode ser utilizada para medir a “qualidade” de um processo de produção com o objetivo de otimizar o processo para que o mesmo seja mais produtivo.

2.4.2 Modelos

Uma parte importante da TSA é a seleção de um modelo probabilístico (ou classe de modelos) adequado para os dados que serão trabalhados (BROCKWELL; DAVIS, 2006). Um modelo de séries temporais para um conjunto de observações x_t é a especificação da distribuição conjunta (ou possivelmente apenas as médias e covariâncias) de uma sequência de variáveis aleatórias X_t das quais x_t é postulada como sendo uma realização (ocorrência).

Modelos auto-regressivos (*Auto-regressive* - AR) são baseados na ideia de que o ponto atual da série, x_t , pode ser explicado através de uma função linear dos p valores passados, x_{t-1} , x_{t-2} ,

..., x_{t-p} , onde p determina o número de passos em direção ao passado necessários para prever o valor atual (SHUMWAY; STOFFER, 2013). Utiliza-se a notação $AR(p)$ para definir um modelo auto-regressivo de ordem p .

Modelos de média móvel (*Moving Average* - MA) por sua vez, utilizam a noção de que as observações da série são uma função de ruído branco (valores independentes com média zero, respeitando uma distribuição normal) composto por dados passados não-observados. Utiliza-se a notação $MA(q)$ para definir um modelo de média móvel de ordem q .

Modelos auto-regressivos de média móvel (*Autoregressive Moving Average* - ARMA) combinam os modelos AR e MA para STs estacionárias, ou seja, séries que não possuem a componente de tendência, de forma que cada ponto observado na ST é uma função de ruído e observações passadas. Utiliza-se a notação $ARMA(p, q)$ para definir um modelo ARMA composto por um modelo AR de ordem p e um modelo MA de ordem q .

Os modelos auto-regressivos integrados de média móvel (*Auto-regressive Integrated Moving Average* - ARIMA) são uma generalização do modelo ARMA para STs não-estacionárias, ou seja, que possuem a componente de tendência. Devem ser executados d passos de diferenciação sobre a série para torná-la estacionária de forma que a série resultante possa ser modelada com um modelo ARMA. Utiliza-se a notação $ARIMA(p, d, q)$ para definir um modelo ARIMA composto por um modelo AR de ordem p , uma diferenciação de grau d e um modelo MA de ordem q .

Após a modelagem de uma ST através dos modelos descritos, pode-se utilizá-los para prever valores futuros. Diversos *software* existentes fornecem funcionalidades para facilitar a modelagem destes modelos. A linguagem de programação R^2 , por exemplo, fornece diversos pacotes para o processamento de computação estatística, incluindo o pacote *forecast* para a previsão de dados futuros utilizando diversos modelos, inclusive os aqui descritos. Essa ferramenta foi utilizada nos experimentos descritos no Capítulo 4.

2.5 Resumo do Capítulo

Neste capítulo foram descritos conceitos base para o entendimento do presente trabalho, incluindo os métodos e técnicas das áreas relacionadas que serviram de base para a elaboração da abordagem proposta neste trabalho.

O presente trabalho utiliza uma rede de colaboração acadêmica que pode ser descrita pelos conceitos da Análise de Redes Sociais. Esta rede de colaboração é utilizada em um Sistema

²The R Project for Statistical Computing: <https://www.r-project.org/>

de Recomendação híbrido, que tem pesquisadores tanto como usuários quanto como itens de recomendação. Por último, a abordagem proposta neste trabalho é a de usar modelos de séries temporais para melhorar as recomendações de colaborações, por isso foram descritos os conceitos básicos da Análise de Séries Temporais.

3 TRABALHOS RELACIONADOS

Neste capítulo serão descritos brevemente alguns trabalhos relacionados à proposta deste trabalho. São trabalhos que abordam métodos de recomendação baseados em fatores temporais, dentro e fora do contexto acadêmico, tendo ou não o foco de colaboração. Primeiramente são mencionados trabalhos com ênfase no fator temporal (Seção 3.1). Em seguida são mencionados os trabalhos com foco em redes acadêmicas e onde já foi explorado o uso de informações temporais neste contexto (Seção 3.2).

3.1 Recomendações Sensíveis ao Tempo

O trabalho “*An MDP-based recommender system*” (SHANI; BRAFMAN; HECKERMAN, 2002) propõe um modelo de SR utilizando os processos de decisão de Markov (*Markov Decision Processes* - MDP), de forma a fugir da visão geralmente utilizada do processo de recomendação como um problema de previsão. Esta abordagem considera o processo de recomendação como um problema de decisão sequencial e leva em consideração os efeitos de longo prazo de cada recomendação e o valor esperado de cada recomendação.

Em “*Temporal diversity in recommender systems*” (LATHIA et al., 2010) os autores fazem um levantamento de algoritmos de recomendação com abordagem colaborativa e observam que o estado da arte inclui algoritmos com baixa diversidade temporal, ou seja, eles recomendam os mesmos itens para os usuários ao longo do tempo. É proposta uma métrica para medir a diversidade temporal e feita uma análise dos fatores que podem influenciar esta diversidade. São projetados e avaliados três métodos para melhorar a diversidade temporal sem penalizar a precisão das recomendações.

Em “*A time-based approach to effective recommender systems using implicit feedback*” (LEE; PARK; PARK, 2008) os pesquisadores apresentam um método para a construção de um SR baseado em filtragem colaborativa efetivo para um ambiente de *e-commerce* sem dados explícitos de *feedback*. Pseudo-dados de avaliações (*ratings*) são construídos a partir do *feedback* implícito utilizando dados temporais (horário em que o usuário comprou um produto ou data de lançamento de um produto). Esses dados são utilizados de forma a valorizar itens comprados recentemente e itens novos no sistema. São utilizados o coeficiente de correlação de Pearson e a similaridade Cosseno para encontrar usuários com preferências similares. Foi realizado um experimento *offline* que mostrou melhores resultados quando comparados com uma abordagem de filtragem colaborativa pura.

Já em “*A recommender system based on tag and time information for social tagging sys-*

tems” (ZHENG; LI, 2011) os autores investigam a importância e utilidade de informações de *tags* e tempo ao prever a preferência de usuários e examinar como explorar tais informações na construção de um modelo de recomendação de recursos. São elaboradas três diferentes abordagens utilizando estas informações e todas mostram resultados satisfatórios. A primeira abordagem supõe que quanto mais uma *tag* é usada pelo usuário, maior interesse ele possui nos recursos marcados com esta *tag*, e que um usuário irá marcar recursos similares com a mesma *tag*. A segunda abordagem supõe que os interesses dos usuários mudam com o tempo e considera as *tags* mais recentes mais relevantes e aplica uma função de esquecimento para as *tags* que não são usadas por muito tempo. A terceira e última abordagem combina as duas anteriores através de uma média ponderada com pesos ajustáveis.

Em “*Hybrid recommender system with temporal information*” (ULLAH et al., 2012) os autores propõem um SR híbrido que leva em conta a similaridade dos atributos dos itens, a similaridade entre avaliações dos usuários, similaridade dos dados demográficos dos usuários e a informação temporal para fazer recomendações. São utilizadas apenas as avaliações feitas pelos usuários em um período recente ao momento das recomendações, o que torna o processo de recomendação muito mais rápido (pois menos informações são consideradas) e as recomendações mais precisas.

No trabalho “*Exploring temporal effects for location recommendation on location-based social networks*” Gao et al. (2013) os pesquisadores apresentam um *framework* de recomendação de localização baseado nas propriedades temporais de movimento do usuário observado em um conjunto de dados reais de redes sociais baseadas em localização. O uso das informações de tempo associadas às localizações do usuário mostrou potencial ao melhorar o desempenho das recomendações de localizações.

Em “*Modeling Short-Term Preferences in Time-Aware Recommender Systems*” (BASILE et al., 2015. Disponível em: <<http://ceur-ws.org/Vol-1388/DeCat2015-paper5.pdf>>. Acesso em: 5 abr. 2016) propõe-se um modelo de SR que utiliza uma função de depreciação e a similaridade do conteúdo de itens computada por modelos semânticos de distribuição para determinar quais avaliações passadas são descartadas com a entrada de novas avaliações. A ideia principal é utilizar informações sobre as preferências dos usuários em curto prazo, pois elas podem mudar mais lenta ou mais rapidamente com o tempo. Os experimentos realizados demonstraram melhora nas recomendações em comparação com uma abordagem de fila simples, onde a primeira avaliação feita é descartada quando uma nova surge.

3.2 Redes Acadêmicas

No trabalho “*Combining social network and semantic concept analysis for personalized academic researcher recommendation*” (XU et al., 2012) é proposta uma rede acadêmica de duas camadas, onde uma delas representa uma rede de colaborações, conectando os pesquisadores, e a outra representa uma rede semântica de conceitos. As duas camadas são conectadas de forma a conectar os pesquisadores aos conceitos em que são especialistas. A rede de colaborações inclui diferentes relações, como por exemplo dois especialistas que participaram de uma mesma conferência, ou foram colegas, ou colaboraram diretamente em algum trabalho publicado. A conexão dos pesquisadores aos conceitos podem ter diferentes pesos, dependendo do grau de atuação do mesmo nas áreas envolvidas. Um estudo foi feito com 35 pesquisadores e a abordagem obteve bons resultados.

Em “Avaliação e Recomendação de Colaborações em Redes Sociais Acadêmicas” (LOPES, 2012) apresenta-se um estudo extensivo a respeito de avaliação de grupos de pesquisadores e recomendação de colaborações em uma rede social acadêmica. É apresentada uma abordagem chamada CORALS (*COLlaboration Recommender for Academic social networks*) que utiliza as métricas de Cooperação, Proximidade Social e Correlação, além de propor métricas sensíveis ao tempo no contexto de recomendações de colaborações acadêmicas pela primeira vez que se tem conhecimento. O presente trabalho apresenta uma abordagem baseada na CORALS e o compara com a mesma.

No trabalho “Recomendação de Colaboração em Redes Sociais Acadêmicas baseada na Afiliação dos Pesquisadores” (BRANDÃO; MORO, 2012), as autoras propõem uma nova metodologia baseada na afiliação institucional dos pesquisadores para recomendar colaborações em uma rede social acadêmica. A métrica proposta *Affin* é calculada similarmente à métrica de Cooperação descrita na Subseção 2.3.2, porém, ao invés de utilizar o número de artigos publicados de um pesquisador i em colaboração com um pesquisador j , é considerado o número de artigos que i publicou em colaboração com pesquisadores da instituição de j . Brandão e Moro (2012) também propõem duas métricas adaptadas para a avaliação do desempenho de SRs de colaborações acadêmicas, sendo elas *Novidade* e *Diversidade*. A Novidade considera recomendações menos populares mais valiosas, enquanto a Diversidade valoriza listas de recomendações com baixa similaridade entre seus itens (baixos valores de Correlação). A avaliação experimental apresentada mostra que a métrica proposta gera recomendações com maior acurácia, novidade e diversidade quando comparada às do estado-da-arte. Descrições mais detalhadas dessas métricas podem ser encontradas na seção 2.2.2.

No trabalho intitulado “*Using link semantics to recommend collaborations in academic social networks*” (BRANDÃO et al., 2013), os autores propõem duas novas métricas para a recomendação de novas colaborações e reforço das colaborações existentes: *Affin*, como descrito em (BRANDÃO; MORO, 2012), e *GLI* (*Geographic Location Information*), que mede as proximidades geográficas entre os pesquisadores da rede acadêmica. Também são propostas três métricas de avaliação: Novidade e Diversidade (descritas em (BRANDÃO; MORO, 2012)), e Cobertura, calculada como uma combinação do índice de Gini com a Entropia de Shannon. As métricas propostas obtiveram resultados melhores que o estado-da-arte quanto à novidade, diversidade e cobertura, tendo *GLI* obtido os melhores resultados. Quanto à precisão, *GLI* apresenta os piores valores para recomendações de iniciação de colaborações, porém mostrou um ótimo impacto nas recomendações de reforço de colaborações. *Affin* apresenta uma melhora quanto à precisão.

Os trabalhos “*ACRec: A Co-authorship based Random Walk Model for Academic Collaboration Recommendation*” (LI et al., 2014) e “*MVCWalker: Random Walk-Based Most Valuable Collaborators Recommendation Exploiting Academic Factors*” (XIA et al., 2014) propõem uma abordagem de recomendação no contexto de coautorias utilizando como base o algoritmo de passeio aleatório com recomeço (*Random Walk with Restart - RWR*). São considerados três fatores: ordem de coautoria, momento da última colaboração e momentos de colaborações. Foram conduzidos experimentos *offline* exaustivos utilizando a base de dados DBLP e foi observada uma melhora significativa nas medidas de precisão, revocação e cobertura, em comparação com o algoritmo RWR.

Em “*Collaborator recommendation in interdisciplinary computer science using degrees of collaborative forces, temporal evolution of research interest, and comparative seniority status*” (CHAIWANAROM; LURSINSAP, 2015), os autores propõem um método de recomendação de colaboração acadêmica que leva em conta 6 fatores: proximidade social, amizade, habilidades complementares, interesses de pesquisa, informações atualizadas e *status* de senioridade. Através de um experimento *offline* foram obtidos resultados superiores com uma diferença estatisticamente significativa em comparação com algoritmos estado-da-arte.

Finalmente, em “*Exploiting Publication Contents and Collaboration Networks for Collaborator Recommendation*” (KONG et al., 2016), os autores propõem uma abordagem chamada CCRec, que utiliza um modelo de agrupamento de tópicos para identificar os domínios acadêmicos e um modelo de passeio aleatório para computar as características dos pesquisadores. Experimentos *offline*, utilizando a base DBLP, demonstram que o CCRec supera os outros algoritmos estado-da-arte em termos de precisão, revocação e medida F1.

3.3 Resumo do Capítulo

Neste capítulo foram descritos diversos trabalhos relacionados com a proposta do presente trabalho. Primeiro foram descritos trabalhos que levam em conta informações temporais de diversas formas para gerar recomendações e depois foram descritos trabalhos no contexto de recomendação de colaborações acadêmicas.

Foi visto que o uso de informações temporais no contexto de recomendações acadêmicas ainda foi pouco explorado e é nesse ponto que a abordagem proposta aqui se encaixa, como descrito no próximo capítulo.

4 RECOMENDAÇÃO DE COLABORAÇÕES ACADÊMICAS UTILIZANDO SÉRIES TEMPORAIS

Na primeira parte deste capítulo serão apresentados a motivação e a visão geral da proposta de utilizar a Análise de Séries Temporais (TSA) para aprimorar a recomendação de colaboração acadêmica (Seção 4.1). Na segunda parte, o uso da Análise de Séries Temporais na abordagem CORALS (LOPES, 2012) é descrito e um exemplo é apresentado para maior compreensão do processo.

4.1 Motivação e visão geral

Como apresentado no Capítulo 3, o uso de informação temporal em SRs foi pouco explorado até o momento. Shani, Brafman e Heckerman (2002), Lee, Park e Park (2008), Zheng e Li (2011), Ullah et al. (2012), Gao et al. (2013) e Basile et al. (2015). Disponível em: <<http://ceur-ws.org/Vol-1388/DeCat2015-paper5.pdf>>. Acesso em: 5 abr. 2016) utilizam informações temporais das ações passadas de um usuário para extrair informações implícitas a respeito de suas preferências. Por exemplo, Basile et al. (2015). Disponível em: <<http://ceur-ws.org/Vol-1388/DeCat2015-paper5.pdf>>. Acesso em: 5 abr. 2016) propõem uma forma de valorizar as preferências dos usuários que foram manifestadas recentemente por eles e ignorar interesses passados que possam ter sido perdidos.

Lathia et al. (2010) citam a importância da diversidade temporal em SRs e apresentam duas perspectivas: mudanças ao longo do tempo e como os usuários respondem às recomendações. É importante lembrar que os SRs tendem a crescer com o tempo, pois novos usuários se juntam ao sistema e novos conteúdos são adicionados, assim como usuários existentes podem atualizar seus perfis, avaliando novos conteúdos. Em uma rede acadêmica, novos pesquisadores começam a publicar trabalhos, ao passo de que novos trabalhos são constantemente publicados, e um pesquisador pode atualizar seu perfil adicionando ou alterando suas áreas de interesse, que podem mudar pouco ou muito ao longo do tempo. Em respeito à forma como os usuários respondem às recomendações, Lathia et al. (2010) descrevem como os usuários reagiram a três diferentes algoritmos de recomendação que utilizaram ao longo do tempo, e concluem que existe um equilíbrio que deve ser encontrado entre a diversidade que os usuários desejam e suas preferências conhecidas pelo sistema. Os usuários demonstraram frustração quando receberam sempre as mesmas recomendações ao longo do tempo e quando receberam recomendações muito diversas, que desconsideraram completamente as preferências dos mesmos.

No contexto de colaborações acadêmicas, Lopes (2012) propôs o uso de dois pesos que levam em conta o aspecto temporal em relação ao passado, considerando as colaborações mais antigas menos relevantes, e valorizando as colaborações mais recentes. Deve-se levar em conta que um pesquisador ativo pode explorar áreas diferentes ao longo do tempo, colaborando com pesquisadores diferentes, e talvez ele não esteja buscando trabalhar novamente com alguma área que não é mais relevante no momento atual. Esse é o primeiro trabalho conhecido que considerou esses fatores temporais no contexto de SRs de colaboração acadêmica. Porém, embora Lopes (2012) considere a importância das alterações da rede acadêmica ao longo do tempo, ela também se limita a utilizar informações temporais passadas.

A proposta do presente trabalho é utilizar as informações temporais disponíveis a respeito do passado de forma a prever as tendências do usuário para o futuro. O método escolhido foi utilizar os conceitos da TSA e o modelo ARIMA, descritos na Seção 2.4, para prever as tendências de publicação e colaboração do usuário nos próximos anos, utilizando como base a abordagem CORALS. A aplicação desse método é descrita na Seção 4.2.

4.2 Aplicação de séries temporais na previsão de métricas

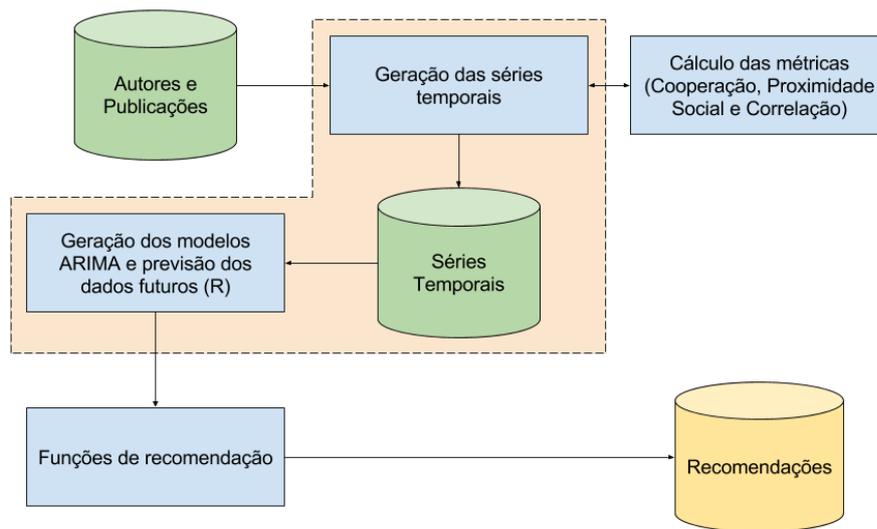
Inicialmente considerou-se a modelagem de STs de trabalhos publicados de cada autor, de forma a tentar prever em quais áreas e com quais colaboradores os seus trabalhos seriam publicados nos próximos anos. A partir dos valores resultantes da previsão, seriam calculadas as métricas de Cooperação, Correlação e Proximidade Social, como na abordagem CORALS, em cima desses novos valores ao invés de utilizar os valores do presente. No entanto, foi constatado que o impacto nas métricas resultantes seria o mesmo do que modelar STs em cima das próprias métricas ao longo do tempo, por isso optou-se por utilizá-las diretamente nas STs.

Para a modelagem das séries foi decidido utilizar todo o histórico de publicações dos autores como base, calculando-se as métricas de Cooperação, Correlação e Proximidade Social para cada ano desde a primeira publicação registrada. Com os modelos definidos, são feitas previsões dos 3 anos seguintes ao período da série, e o valor encontrado para o terceiro ano é utilizado como a nova métrica. Desta forma, são propostas três novas métricas: Cooperação TS, Proximidade Social TS e Correlação TS.

Primeiramente, são geradas séries de observações para as métricas de Cooperação, Proximidade Social e Correlação para cada par de autores a cada ano. Em seguida, um modelo ARIMA é ajustado a partir destas observações e então são geradas as previsões para os próximos três anos para cada métrica, para cada par de autores. Os valores resultantes para o terceiro ano das

previsões são utilizados pelas funções de recomendação em substituição dos valores convencionais das métricas da abordagem CORALS. As taxas de erro nas previsões são ignoradas e são utilizadas as médias das previsões. Por fim, as recomendações são geradas. A figura 4.1 ilustra este processo.

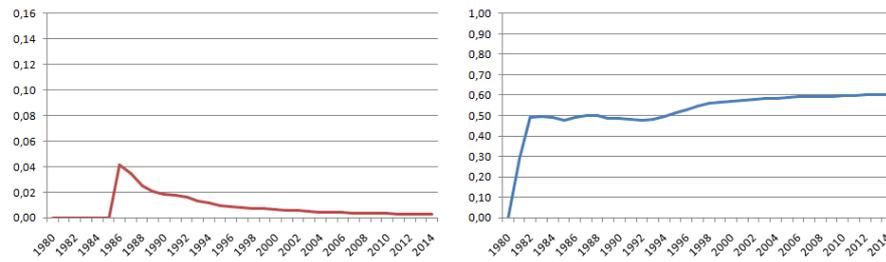
Figura 4.1 – Processo de recomendação utilizando a abordagem proposta, destacada pelo contorno tracejado.



Fonte: o autor

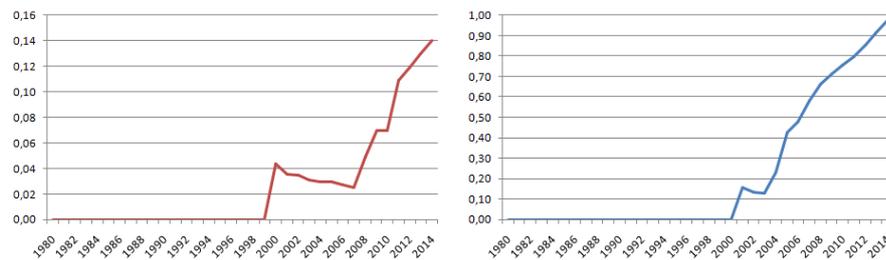
As figuras 4.2 e 4.3 mostram exemplos de séries temporais geradas a partir de dois pares de autores distintos. No primeiro caso, é possível observar que os autores *A* e *B* publicaram trabalhos conjuntos na década de 80, porém outras colaborações ocorreram nos anos seguintes e esta colaboração não voltou a acontecer, de forma que a medida de cooperação entre os dois foi diminuindo com o tempo. Porém, também é possível observar que a correlação entre as áreas de atuação dos dois foi mediana durante todo o período considerado, sendo levemente elevada a partir da década de 90. No segundo caso, os autores *C* e *D* trabalharam juntos pela primeira vez nos anos 2000, ficaram alguns anos sem colaborar um com o outro, e trabalharam juntos novamente em 2010 e nos anos seguintes. Também pode-se observar, a partir da correlação entre os dois, que ambos trabalharam cada vez mais em áreas similares, possivelmente convergindo para alguma área em comum, o que faz sentido considerando que a cooperação entre eles é crescente. Nas quatro STs apresentadas, os últimos três anos (2012, 2013 e 2014) representam os valores previstos utilizando o modelo ARIMA, e os valores respectivos ao ano 2014 são os valores utilizados pelas funções de recomendação.

Figura 4.2 – Exemplo 1: STs de Cooperação (esquerda) e Correlação (direita) entre os autores *A* e *B*.



Fonte: o autor

Figura 4.3 – Exemplo 2: STs de Cooperação (esquerda) e Correlação (direita) entre os autores *C* e *D*.



Fonte: o autor

4.3 Resumo do Capítulo

Neste capítulo foram apresentadas as motivações e visão geral da abordagem recomendada pelo presente trabalho. Foram descritos os passos utilizados pelo SR proposto e a abordagem específica com o uso da Análise de Séries Temporais. Por último, foram apresentados exemplos de STs geradas para a determinação das métricas de Cooperação e Correlação para dois pares de autores, de forma a ilustrar o papel dos modelos ARIMA no processo.

5 EXPERIMENTOS

No capítulo anterior foram apresentados a motivação e a visão geral sobre a proposta de utilizar a Análise de Séries Temporais (TSA) para aprimorar a recomendação de colaborações acadêmicas. Neste capítulo serão descritos os experimentos práticos executados como forma de avaliar se a utilização da TSA melhorou ou não o desempenho da abordagem CORALS (LOPES, 2012).

Na primeira parte, será descrito como os dados utilizados foram obtidos (Seção 5.1) e como as listas de recomendações foram geradas (Seção 5.2). Na segunda parte, são descritas as duas avaliações realizadas para comparar a abordagem modificada e a abordagem CORALS (LOPES, 2012) utilizada como base e o desempenho da abordagem resultante em relação aos usuários. Primeiro, é descrita a avaliação experimental *offline* que foi feita separando os dados obtidos em dois momentos e utilizando os dados mais antigos como base para o modelo de recomendação e os dados mais recentes como conjunto-verdade (i.e., *gold standard*) para avaliação da qualidade das recomendações produzidas (Subseção 5.3.1). Depois, é descrito o estudo de usuários, que utilizou dados obtidos de pesquisadores aos quais foram apresentadas recomendações obtidas com o método proposto (Subseção 5.3.2). Por último, é feita uma análise geral dos resultados obtidos (Subseção 5.3.3).

5.1 Fonte de dados

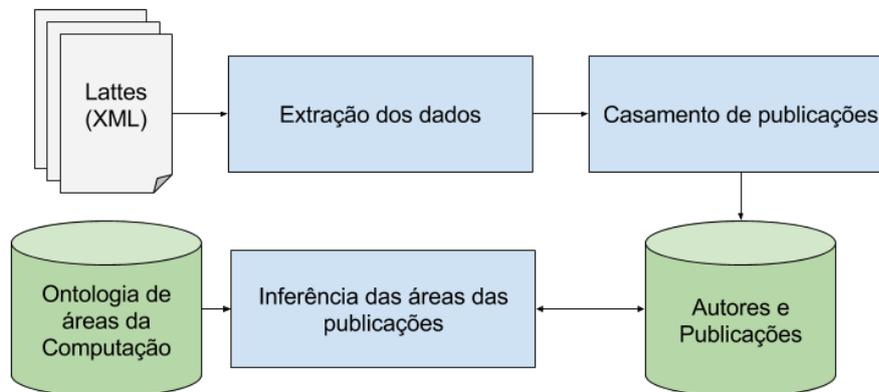
Como base para os experimentos, foi escolhido o corpo docente do Instituto de Informática da UFRGS, onde há uma diversidade grande de pesquisadores nas mais diversas áreas da Computação. Os dados das publicações dos pesquisadores foram obtidos pela Plataforma Lattes¹, através dos currículos Lattes dos pesquisadores. Foram selecionados 50 pesquisadores do Instituto de Informática da UFRGS que publicaram trabalhos regularmente nos últimos anos, em sua grande maioria, professores que atuam na pós-graduação. A Figura 5.1 ilustra o processo de extração e organização dos dados, que é explicado nas próximas subseções.

5.1.1 Currículo Lattes

A partir dos currículos dos pesquisadores, foram extraídos seus nomes completos e uma lista de publicações. Para cada publicação foram extraídos o título completo e o ano de publicação. As áreas das publicações são preenchidas livremente pelos pesquisadores, sem um

¹Plataforma Lattes: <http://lattes.cnpq.br/>

Figura 5.1 – Processo de extração e organização dos dados dos pesquisadores e publicações.



Fonte: o autor

padrão definido, portanto foram descartadas. Foi decidido descartá-las pois a utilização delas traria muitas áreas diferentes e o casamento de autores por áreas de interesse seria prejudicado.

Os currículos foram obtidos manualmente no site da Plataforma Lattes², em formato XML (*Extensible Markup Language*). Para a extração dos dados foi utilizada uma ferramenta própria, desenvolvida pelo autor, implementada em Java e utilizando a biblioteca JSoup³, que permite a fácil extração de dados em HTML e XML. As localizações dos dados extraídos dos XMLs dos currículos são listadas a seguir, no formato XPath⁴:

- Nome completo do pesquisador:
`/curriculo-vitae/dados-gerais@nome-completo`
- Título da publicação (artigos publicados):
`/curriculo-vitae/producao-bibliografica/artigos-publicados`
`/artigo-publicado/dados-basicos-do-artigo@titulo-do-artigo`
- Ano da publicação (artigos publicados):
`/curriculo-vitae/producao-bibliografica/artigos-publicados`
`/artigo-publicado/dados-basicos-do-artigo@ano-do-artigo`
- Título da publicação (trabalhos em eventos):
`/curriculo-vitae/producao-bibliografica/trabalhos-em-eventos/trabalho-em-eventos/dados-basicos-do-trabalho`
`@titulo-do-trabalho`
- Ano da publicação (trabalhos em eventos):

²Buscar Currículo Lattes: <http://buscatextual.cnpq.br/buscatextual/busca.do?metodo=apresentar>

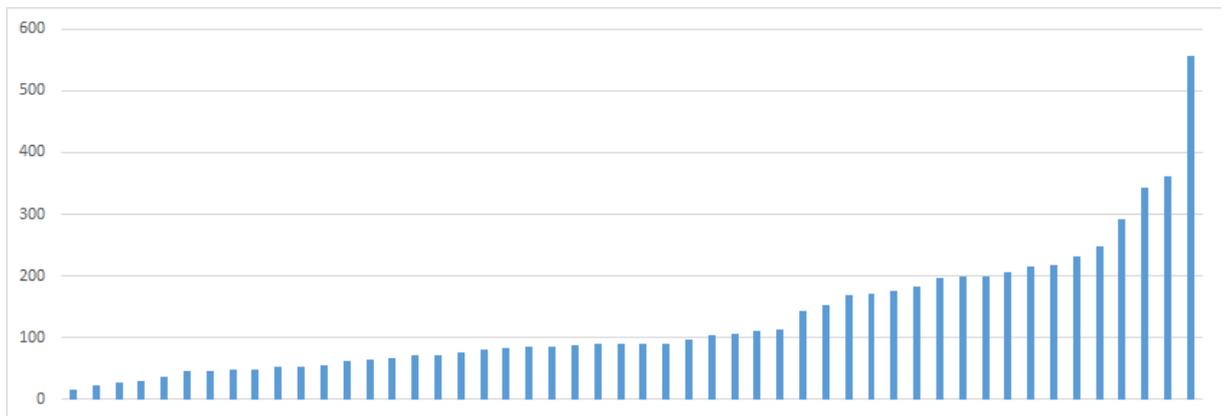
³JSoup: <http://jsoup.org/>

⁴XML Path Language (XPath): <https://www.w3.org/TR/xpath/>

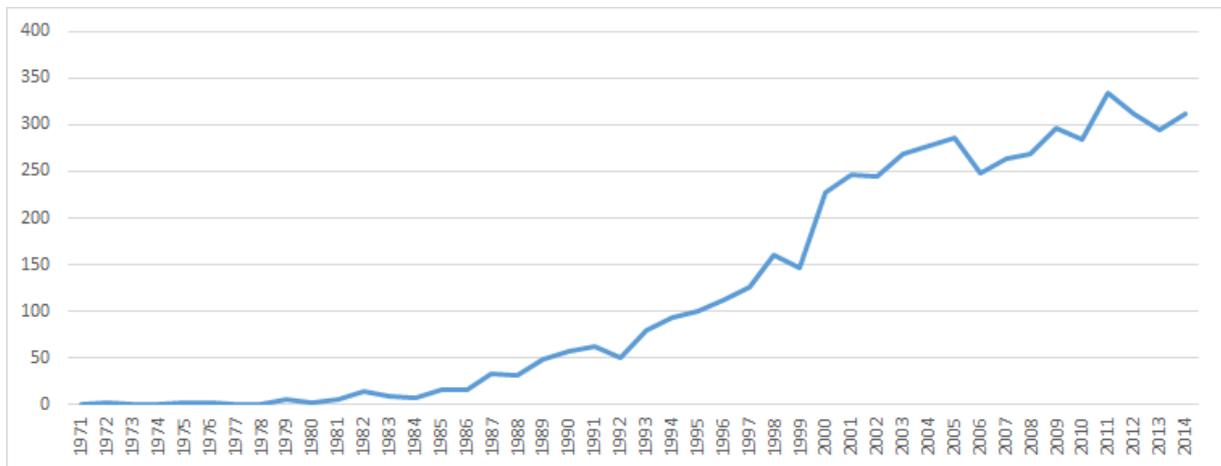
/curriculo-vitae/producao-bibliografica/trabalhos-em-
eventos/trabalho-em-eventos/dados-basicos-do-trabalho
@ano-do-trabalho

A Figura 5.2 apresenta um gráfico com a quantidade de publicações de cada pesquisador (a) e um gráfico com a quantidade de publicações por ano de publicação (b).

Figura 5.2 – Gráfico mostrando a quantidade de publicações de cada autor (a) e por ano (b) da base de dados utilizada nos experimentos.



(a) Por autor



(b) Por ano

Fonte: o autor

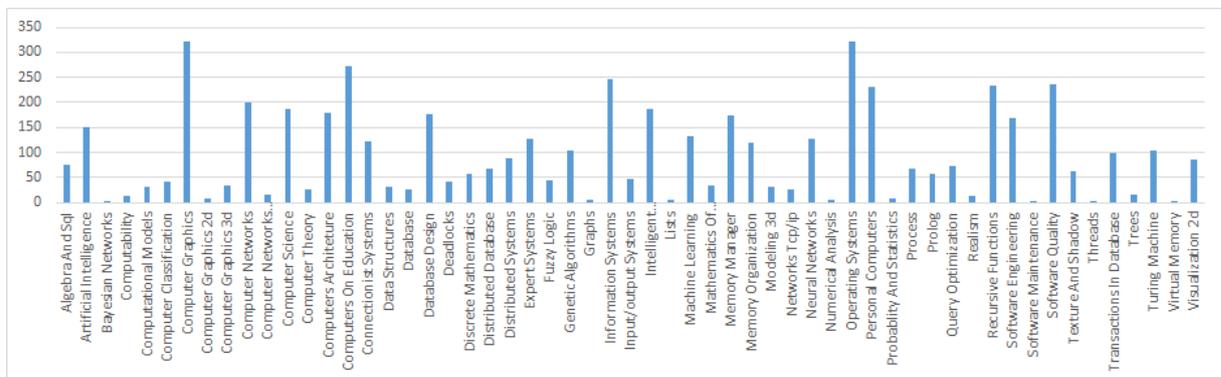
5.1.2 Definição das áreas das publicações

O currículo Lattes é preenchido livremente por cada pesquisador, por isso corre o risco de ter alguns dados faltando ou preenchidos com informações sem um padrão bem definido. Um dos casos observados foi a área de pesquisa das publicações, que foi preenchida com sistemas

de classificação arbitrários, ou sem um sistema de classificação, o que tornaria difícil uma classificação geral das publicações.

Para resolver esse problema, foi decidido utilizar a ontologia de áreas de pesquisa em computação proposta por Loh et al. (2006) e utilizada por Lopes et al. (2010). Essa ontologia utiliza uma classificação das áreas da Ciência da Computação similar à versão de 1998 do sistema de classificação da ACM (*Association for Computing Machinery*)⁵. Embora a ACM tenha publicado um novo sistema de classificação em 2012⁶ que substitui a versão de 1998 e inclui áreas que surgiram mais recentemente, optou-se por utilizar a ontologia existente, pois a criação de uma nova ontologia para classificação está fora do escopo deste trabalho. A Tabela 5.1 mostra todas as áreas da ontologia utilizada e a quantidade de publicações que foi classificada para cada uma delas. A Figura 5.3 mostra um gráfico comparativo das quantidades de publicações classificadas para cada área. Para definir as áreas das publicações dentro dessa ontologia é utilizado um dicionário que associa conceitos (ou termos) às diferentes áreas. Os termos contidos nos títulos das publicações são extraídos e utilizados para calcular as pontuações da publicação para cada área da ontologia. A área que obtiver a maior pontuação é selecionada como a área em que a publicação melhor se adequa.

Figura 5.3 – Gráfico mostrando a quantidade de publicações por área da base de dados utilizada nos experimentos.



Fonte: o autor

5.1.3 Casamento de coautorias das publicações

Assim como o caso das áreas de pesquisa, observou-se inconsistência nas coautorias das publicações cadastradas nos Lattes dos pesquisadores. Isso acontece porque cada pesquisador

⁵The 1998 ACM Computing Classification System: <http://www.acm.org/about/class/ccs98-html>

⁶The 2012 ACM Computing Classification System: <http://www.acm.org/about/class/2012>

Tabela 5.1 – Áreas de pesquisa e a quantidade de publicações para cada uma.

Área de pesquisa	Qtd. de publicações
Algebra And Sql	76
Artificial Intelligence	150
Bayesian Networks	1
Computability	12
Computational Models	32
Computer Classification	42
Computer Graphics	322
Computer Graphics 2d	8
Computer Graphics 3d	33
Computer Networks	200
Computer Networks Architecture	15
Computer Science	187
Computer Theory	26
Computers Architeture	180
Computers On Education	273
Connectionist Systems	121
Data Structures	32
Database	25
Database Design	177
Deadlocks	42
Discrete Mathematics	58
Distributed Database	68
Distributed Systems	89
Expert Systems	127
Fuzzy Logic	45
Genetic Algorithms	104
Graphs	6
Information Systems	246
Input/output Systems	46
Intelligent Agents/multiagent Systems	186
Lists	5
Machine Learning	132
Mathematics Of Computing	35
Memory Manager	174
Memory Organization	120
Modeling 3d	30
Networks Tcp/ip	27
Neural Networks	127
Numerical Analysis	4
Operating Systems	321
Personal Computers	230
Probablity And Statistics	9
Process	67
Prolog	58
Query Optimization	73
Realism	13
Recursive Functions	234
Software Engineering	169
Software Maintenance	2
Software Quality	235
Texture And Shadow	62
Threads	3
Transactions In Database	99
Trees	15
Turing Machine	104
Virtual Memory	1
Visualization 2d	85

Fonte: o autor

é responsável por cadastrar os dados de suas publicações e não há um cadastro geral de publicações, o que acarretou em uma mesma publicação ter pequenas variações em seus títulos e nos nomes dos coautores em currículos diferentes.

Para resolver esse problema, os títulos das publicações foram casados (comparados), transformando todo texto em letras minúsculas e aplicando-se o algoritmo de “Distância de Edição” para calcular a similaridade entre eles. A Distância de Edição entre duas sequências de caracteres x e y é definida como o custo mínimo de uma sequência de operações que transformam x em y (NAVARRO, 2001). O custo dessa sequência de operações é o somatório do custo de cada operação. Esse algoritmo permite o casamento aproximado, ou seja, permite erros até um determinado limiar que deve ser determinado para cada aplicação, de duas sequências de caracteres. Mais detalhes podem ser encontrados em (NAVARRO, 2001) e (RISTAD; YIANILOS, 1998).

A similaridade entre os títulos das publicações é definida pelo custo calculado pela Distância de Edição dividido pelo tamanho do maior título. Foi utilizado um limiar de 70% de similaridade definido através de uma busca binária manual sobre os dados extraídos dos currículos, em que se reduziu o número de casamentos incorretos sem perder os casamentos corretos.

Também foram observadas publicações com títulos similares feitas pelos mesmos pesquisadores em anos consecutivos. Por isso, também foram utilizados os anos das publicações como critério para o casamento. Diante disso, uma publicação só foi considerada igual a outra com título similar caso tivesse sido publicada no mesmo ano.

5.2 Geração de recomendações

Para a geração das recomendações foi desenvolvida uma ferramenta utilizando as linguagens de programação Java⁷ e R⁸. A linguagem R foi utilizada para a geração dos modelos ARIMA e previsão dos valores futuros das séries temporais, enquanto que a linguagem Java foi utilizada para todas as outras partes do processo.

Os dados dos pesquisadores e suas publicações são carregados e a partir deles são calculadas as métricas de Cooperação, Proximidade Social e Correlação, descritas na Subseção 2.3.2. Depois, os pesos com fator temporal Cooperação Tr e Cooperação Tg são calculados junto com os respectivos valores de Proximidade Social utilizando seus resultados. Estes dois pesos foram calculados utilizando o valor 0 para a variável t_{min} . Por último, são calculadas as métricas propostas utilizando o modelo ARIMA, como descrito na Seção 4.2.

⁷Java: <http://www.oracle.com/br/java/overview/index.html>

⁸The R Project for Statistical Computing: <https://www.r-project.org/>

A modelagem das STs é feita através da linguagem de programação *R* e o pacote *forecast* disponível para a mesma. São utilizadas as funções “*auto.arima*”, para gerar os modelos ARIMA das séries utilizadas, e “*forecast.Arima*” para calcular os valores futuros das séries. A função “*auto.arima*” busca o melhor modelo ARIMA possível sem necessidade de definir os parâmetros p , d e q , já que a própria função encontra os valores ideais para estes parâmetros. A função “*forecast.Arima*”, por sua vez, calcula os valores futuros da ST utilizando o modelo encontrado.

A última etapa do algoritmo é gerar duas listas de recomendação para cada pesquisador: uma lista de colaborações existentes, que podem ser reforçadas em busca de bons resultados, e outra lista com colaborações que têm um bom potencial e podem ser iniciadas. São utilizadas as funções descritas na Subseção 2.3.3 para gerar os *rankings*, a partir das diferentes combinações de métricas descritas a seguir:

- Configuração 1 (Base): Cooperação, Correlação e Proximidade Social;
- Configuração 2 (TR): Cooperação (com peso Tr), Correlação e Proximidade Social;
- Configuração 3 (TG): Cooperação (com peso Tg), Correlação e Proximidade Social;
- Configuração 4 (TS): Cooperação TS, Correlação TS e Proximidade Social TS.

Para o cálculo dos *rankings* de recomendações para novas colaborações foram utilizados os valores 1 e 150 para os pesos w_{Cr} e w_{Sc} , respectivamente (ver Subseção 2.3.2). Estes valores foram os que apresentaram melhores resultados no trabalho de Lopes (2012) e por esse motivo foram utilizados no presente trabalho.

Após a geração das recomendações, um usuário poderia avaliar as recomendações como forma de *feedback* para o sistema, permitindo que o mesmo tenha mais informações para gerar recomendações no futuro. Bostandjiev, O’Donovan e Höllerer (2012) mostraram que a interação do usuário pode melhorar não só a precisão das recomendações futuras, mas também a experiência do mesmo com o sistema. Porém, o presente trabalho limita-se a avaliar comparativamente os algoritmos utilizados sem incluir informações externas aos currículos dos pesquisadores.

5.3 Avaliações experimentais

Nesta seção são descritas as avaliações experimentais executadas e seus resultados.

5.3.1 Avaliação *offline*

A avaliação experimental *offline* foi feita separando-se os dados em dois conjuntos, representando dois períodos disjuntos. O primeiro conjunto foi utilizado pelos algoritmos de recomendação para gerar as recomendações, por isso foi chamado de Conjunto Base. O segundo conjunto foi utilizado para validar as recomendações feitas, executando o papel de *Ground Truth* (ou *Gold Standard*), por isso foi chamado de Conjunto-Verdade. Todos os dados de publicações até o ano 2011 foram utilizados como Conjunto-Base e os dados de publicações de 2012 à 2014 como Conjunto-Verdade, de forma a verificar se, utilizando apenas os dados até 2011, seria possível prever as colaborações que ocorreram ou foram reforçadas a partir de 2012.

Após as etapas de extração dos dados, cálculo das métricas e definição dos *rankings*, os *rankings* gerados foram avaliados através das métricas de precisão, revocação, precisão até 10 (Pr@10) e média das precisões médias (MAP), descritas na Subseção 2.2.2. Após o cálculo das métricas, os resultados foram comparados utilizando o Teste-T bicaudal com p-valor 5% para determinar se a diferença entre eles é estatisticamente significativa.

Os resultados podem ser observados nas Tabelas 5.2 (para recomendações de reforço das colaborações) e 5.3 (para recomendações de iniciação de colaborações). Como pode ser visto na Tabela 5.2, a abordagem proposta (TS) obteve resultados inferiores em relação às outras abordagens, na maioria dos casos. Na Tabela 5.3, porém, ela foi superior às abordagens Base e TG, mas levemente inferior à abordagem TR.

Tabela 5.2 – Médias dos valores das métricas de avaliação para cada abordagem em relação à recomendação de colaborações que devem ser reforçadas.

	Precisão	Revocação	P@10	MAP
Base	25,40%	64,00%	10,80%	27,00%
TR	42,67%	49,77%	8,40%	41,29%
TG	25,77%	64,00%	10,80%	27,16%
TS	21,70%	52,77%	8,80%	21,19%

Fonte: o autor

Nas Tabelas 5.4 e 5.5 podem ser vistos os valores de Teste-T calculados para comparação das abordagens com a abordagem proposta. A maioria dos resultados para reforço de recomendações apresenta diferenças estatisticamente significativas, enquanto os resultados para iniciação de colaborações apresentaram valores muito similares e sem diferenças estatisticamente significativas.

Também foram calculadas as métricas de Cobertura de Catálogo, Novidade e Diversidade,

Tabela 5.3 – Médias dos valores das métricas de avaliação para cada abordagem em relação à recomendação de colaborações que devem ser iniciadas.

	Precisão	Revocação	P@10	MAP
Base	0,00%	0,00%	2,00%	3,40%
TR	0,53%	14,23%	2,40%	3,97%
TG	0,00%	0,00%	2,00%	3,40%
TS	0,47%	11,23%	2,40%	3,91%

Fonte: o autor

Tabela 5.4 – Resultados dos Testes-T executados para comparar a abordagem proposta com os outros algoritmos para recomendações de reforço de colaborações.

	Base	TR	TG
Precisão	16,17%	0,00%	12,52%
Revocação	0,73%	18,24%	0,73%
Precisão@10	0,29%	15,94%	0,29%
MAP	4,35%	0,00%	3,79%

Fonte: o autor

propostas por Ge, Delgado-Battenfeld e Jannach (2010) e Brandão e Moro (2012), que podem ser vistas nas Tabelas 5.6, para recomendações de reforço de colaboração, e 5.7, para recomendações de iniciação de colaborações.

A Cobertura de Catálogo dos algoritmos é similar tanto nas recomendações de reforço de colaborações quanto nas de iniciação, apresentando valores muito próximos de 100%. Já as Novidades apresentaram valores próximos de 100% para as recomendações de reforço de colaborações, enquanto nas recomendações de iniciação apresentaram valores mais próximos de zero. Acredita-se que isso ocorreu pois o grupo de pesquisadores utilizado é bem conectado, já que todos atuam na mesma instituição, o que acarretou em recomendações muito similares entre todos os pesquisadores. Por último, a Diversidade apresentou valores similares para todos os algoritmos e tipos de recomendações, com uma média aproximada de 65%. No geral todas as abordagens tiveram um desempenho similar em relação a essas métricas, o que era esperado considerando que todas elas utilizam mecanismos comuns para gerar as recomendações.

Quanto à serendipidade dos algoritmos, não há uma forma bem estabelecida de medir esse valor, mas é geralmente aceito que ele é dependente de fatos subjetivos, como o sentimento do usuário ao receber a recomendação. Como neste experimento não é possível prever o comportamento do usuário ao receber as recomendações, não foi possível determinar a serendipidade dos algoritmos. Porém foi observado que muitas recomendações não se tornaram colaborações concretas, especialmente as recomendações de iniciação de colaboração, o que pode significar

Tabela 5.5 – Resultados dos Testes-T executados para comparar a abordagem proposta com os outros algoritmos para recomendações de iniciação de colaborações.

	Base	TR	TG
Precisão	0,29%	34,45%	0,29%
Revocação	0,73%	18,24%	0,73%
Precisão@10	15,94%	0,00%	15,94%
MAP	10,46%	57,84%	10,42%

Fonte: o autor

Tabela 5.6 – Valores das métricas Cobertura de Catálogo, Novidade e Diversidade para cada abordagem em relação à recomendação de colaborações que devem ser reforçadas.

	Cobertura	Novidade	Diversidade
Base	90,00%	92,00%	63,01%
TR	76,00%	96,00%	78,39%
TG	90,00%	92,00%	62,91%
TS	88,00%	96,00%	62,40%

Fonte: o autor

que os algoritmos possuem uma boa serendipidade e os pesquisadores poderiam se beneficiar das recomendações para encontrar colaborações que de outra forma não encontrariam. Para determinar se isso ocorreria realmente, devem ser feitos estudos de usuários e experimentos *online* (sugestão para trabalhos futuros).

O método de experimentação *offline* permite verificar o potencial dos algoritmos de recomendação e compará-los em certo nível. Porém, os méritos dos algoritmos só podem ser medidos completamente através do conhecimento das ações do usuário após as recomendações. O usuário já tinha conhecimento sobre os colaboradores recomendados? O usuário não entendeu o motivo ou discordou das recomendações? A produtividade dos grupos de pesquisa aumentou a partir da introdução ou reforço das colaborações recomendadas? De forma a complementar os resultados, devem ser feitos estudos de usuários ou experimentos *online* com usuários reais que possam fornecer *feedback*. Neste trabalho propõe-se um estudo de usuários com um grupo de 16 pesquisadores, que é descrito na próxima seção.

5.3.2 Estudo de usuários

Para este experimento as recomendações foram geradas de forma idêntica às do experimento *offline*, porém foram utilizados todos os dados das publicações até o ano 2014, enquanto que no experimento anterior foram considerados apenas os dados até o ano 2011. Foram utilizados

Tabela 5.7 – Valores das métricas Cobertura de Catálogo, Novidade e Diversidade para cada abordagem em relação à recomendação de colaborações que devem ser iniciadas.

	Cobertura	Novidade	Diversidade
Base	98,00%	12,00%	64,60%
TR	98,00%	8,00%	64,32%
TG	98,00%	12,00%	64,58%
TS	98,00%	12,00%	65,06%

Fonte: o autor

os mesmos dados da extração previamente descrita, porém apenas as métricas que utilizam a TSA foram consideradas. Por fim, foram definidos os *rankings* e selecionados os três primeiros colocados de cada um para entrar na pesquisa.

A avaliação deste experimento foi constituída por uma série de formulários a serem preenchidos por pesquisadores voluntários a respeito de recomendações feitas a eles. Os 16 pesquisadores participantes receberam seis recomendações ao todo: três recomendações de colaborações existentes, que poderiam produzir bons resultados sendo reforçadas, e três recomendações de novas colaborações, que poderiam trazer bons frutos. Para cada recomendação, foram feitas as seguintes perguntas objetivas:

- Você conhece ou já ouviu falar de *X*? (Sim, Não)
- Você já leu algum trabalho de *X*? (Sim, Não, Não lembro)
- Você já trabalhou com *X*? (Sim, Não)
 - Caso não tenha trabalhado, você trabalharia com *X*? (Sim, Não, Talvez)
- Na sua opinião, há chances de que *X* tenha publicado algum trabalho recentemente nas seguintes áreas?
 - Nesta pergunta foram apresentadas as três áreas em que o pesquisador mais publicou em toda sua carreira acadêmica;
 - Para cada área apresentada, as respostas possíveis eram: Sim, Não e Talvez.
- Além das questões objetivas, foi dada a opção para comentários em texto livre, caso o voluntário tivesse algo a acrescentar.

As três áreas em que o pesquisador recomendado mais publicou são apresentadas para que, além dos resultados mais básicos da recomendação, o usuário do experimento pudesse dar indicadores quanto à capacidade do sistema desenvolvido fazer recomendações. Como as áreas das publicações foram definidas automaticamente através de uma ontologia, é possível que haja

uma certa imprecisão nessas atribuições, e como as áreas são um elemento de grande impacto nos algoritmos utilizados, é possível que ocorram recomendações incorretas por causa de publicações classificadas incorretamente. Dessa forma, caso o usuário discorde das áreas apontadas pelo sistema, pode-se extrair mais informações do restante de suas respostas. Se o colaborador recomendado não atua em áreas de interesse do usuário, faz sentido que a recomendação seja desconsiderada.

Inicialmente foram observados alguns dados básicos a partir das respostas dos participantes do estudo:

- Todos os usuários já conheciam os 6 colaboradores recomendados para si;
- 45,8% dos colaboradores recomendados para iniciação de colaboração já haviam trabalhado com o usuário;
- Dos 54,2% restantes, apenas 1 das recomendações foi descartada, com 18,75% das recomendações sendo completamente aceitas pelos usuários (resposta “sim” para se trabalharia com o pesquisador recomendado) e 33,33% das recomendações sendo parcialmente aceitas (resposta “talvez”);
- Em torno de 87,5% dos colaboradores recomendados para reforço de colaborações já haviam de fato trabalhado com o usuário;
- Dos 12,5% restantes, 50% das recomendações foram aceitas completamente (resposta “sim”) e os outros 50% foram aceitas parcialmente (resposta “talvez”);
- Os usuários tiveram dúvida em pelo menos uma das áreas listadas para 56,25% dos pesquisadores recomendados para iniciação de colaboração e 60,4% dos pesquisadores recomendados para reforço de colaboração;
- Os usuários concordaram com apenas uma das áreas listadas em 43,75% das recomendações de iniciação de colaboração e 29,17% das recomendações de reforço de colaboração;
- Os usuários discordaram das três áreas listadas em 22,9% das recomendações de iniciação de colaboração e 20,8% das recomendações de reforço de colaboração.

Por se tratar de pesquisadores que atuam na mesma instituição de pesquisa, era esperado que a maioria dos usuários já estivessem familiarizado com os pesquisadores recomendados. Da mesma forma, como as questões foram formuladas no sentido de “trabalhar juntos”, e não de “publicar algum trabalho juntos”, era esperado que, mesmo com a ausência de publicações em colaboração na base de dados utilizada, algumas colaborações já tivessem ocorrido. Também é possível que os dados utilizados estivessem desatualizados e portanto carecessem de informações sobre algumas publicações. Já o caso das colaborações existentes que deveriam

ser intensificadas, mas foram identificadas como inexistentes pelos usuários, pode ser explicado pela natureza de determinados trabalhos em que os autores colaboram de alguma forma sem interagir com todos os colaboradores, ou apenas foi uma colaboração pontual e o pesquisador não considerou como uma colaboração efetiva. Porém, isso não afeta o propósito do algoritmo, que é identificar colaborações que podiam ser mais fortes considerando as áreas em que os pesquisadores atuam, e pode-se dizer a reação dos usuários a essas recomendações foi positiva.

Porém, mesmo com resultados satisfatórios, com apenas uma recomendação descartada, a definição das áreas de pesquisa de cada pesquisador recomendado foi negativamente avaliada em geral. Isso pode ter ocorrido porque as áreas das publicações foram definidas com base apenas nas palavras dos títulos das mesmas, pela ontologia usada possuir conceitos não mais utilizados (é similar à classificação da ACM de 1998) ou ambos.

5.3.3 Análise dos resultados

O experimento *offline* apresentou valores baixos nas métricas de avaliação provenientes da RI, havendo uma piora estatisticamente significativa em relação às outras abordagens nas recomendações de reforço de colaborações. No caso das recomendações de iniciação de colaboração, as abordagens Base e TG obtiveram valores muito inferiores, porém a abordagem proposta não teve diferença estatisticamente significativa em relação à TR, tendo ambas obtido valores baixos. Isso provavelmente ocorreu porque não se levou em conta o *feedback* e o comportamento dos usuários frente às recomendações. Em vez disso, apenas as colaborações que surgiram no período avaliado (2012-2014) é que foram consideradas relevantes. De acordo com Lopes (2012), isso mostra que um SR pode ser útil para essa rede de colaborações, pois, caso os valores de precisão fossem altos, significaria que a rede já está se comportando de maneira a otimizar seu potencial de colaboração.

Já no caso das métricas de Cobertura, Novidade e Diversidade, todas as abordagens obtiveram resultados similares, concluindo-se que o uso de séries temporais não agregou valor aos algoritmos usados.

Quanto ao estudo de usuários, a grande maioria das recomendações foi avaliada positivamente, porém a classificação de áreas das publicações foi avaliada negativamente, com uma taxa de erros grande. Isso pode ter ocorrido porque essa classificação foi feita com base apenas nos títulos dos trabalhos, o que muitas vezes pode não ser o suficiente. Alguns pesquisadores procuram atrair a atenção de possíveis leitores colocando títulos chamativos em seus trabalhos,

mas que não revelem tanto a respeito da área em que o mesmo se enquadra. Usar os resumos dos trabalhos para determinar a área em que se enquadram seria o ideal, porém nem sempre se tem acesso a essas informações, como foi o caso deste trabalho. Também se acredita que a ontologia utilizada para classificação não possua em seu dicionário de conceitos alguns termos originados nos últimos anos, impedindo que os trabalhos mais recentes fossem classificados apropriadamente. Considerado isso, acredita-se que a métrica de Correlação não tenha influenciado tanto nas recomendações, pois, nesse caso, mais recomendações teriam sido descartadas.

Por outro lado, o conjunto de dados utilizado continha apenas pesquisadores de uma única instituição, o que limitou bastante o experimento. Em uma mesma instituição, mesmo em áreas de pesquisa completamente diferentes, a chance de dois pesquisadores não se conhecerem é baixa, e através do convívio no espaço físico da instituição é possível que se crie simpatia entre eles, o que diminuiria as chances de um pesquisador recusar a recomendação para trabalhar com um colega.

Além disso, seria necessário um experimento mais completo para avaliar completamente o impacto do SR nas pesquisas desenvolvidas pelos usuários, sendo observados o comportamento de dois grupos distintos durante um período mais longo, em que um grupo receberia as recomendações e o outro não, de forma a comparar o comportamento dos dois grupos após o momento das recomendações.

5.4 Resumo do Capítulo

Neste capítulo foram descritos todos os passos utilizados pelo SR proposto e a abordagem específica com o uso da Análise de Séries Temporais. Por último, foram descritas as avaliações feitas através dos experimentos e as análises dos resultados obtidos.

6 CONCLUSÃO

O presente trabalho propôs uma nova abordagem de recomendação sensível ao tempo através do uso da análise de séries temporais com a finalidade de melhorar as recomendações de colaborações no contexto acadêmico. Para avaliação da abordagem proposta, foram feitos dois experimentos: um experimento *offline* e um estudo de usuários. Para ambos os experimentos implementou-se um SR utilizando dados de currículos Lattes de 50 pesquisadores do Instituto de Informática da UFRGS e foram geradas recomendações de novas colaborações e colaborações existentes que poderiam ser reforçadas para cada pesquisador. No experimento *offline*, os três últimos anos do período abrangido pelos dados utilizados (até 2014) foram separados em um conjunto-verdade (*gold standard*), e o restante dos dados foi utilizado para gerar recomendações. As recomendações geradas foram comparadas com as colaborações do conjunto-verdade a fim de verificar a precisão das recomendações. Por último, 16 dos 50 pesquisadores participaram de um estudo de usuários onde cada um recebeu 3 recomendações de colaborações existentes que poderiam ser reforçadas e 3 recomendações de novas colaborações, e preencheram um formulário padrão para cada recomendação.

No experimento *offline*, a abordagem proposta foi comparada com métodos de estado-da-arte em recomendações no contexto de colaborações acadêmicas e os resultados foram analisados através de métricas tradicionais de RI. O desempenho da abordagem proposta teve uma piora estatisticamente significativa em comparação com os outros algoritmos no caso das recomendações de reforço de colaborações, e resultados similares a uma das abordagens e significativamente superiores às outras duas, no caso de recomendações de iniciação de colaborações. É de conhecimento que as métricas utilizadas têm demonstrado resultados abaixo do esperado em diversos trabalhos sobre SRs e a área em geral está no caminho de estabelecer métricas mais adequadas para este tipo de avaliação. Por conta disto, foram utilizadas também mais três métricas que estão sendo utilizadas em trabalhos mais recentes: Cobertura, Diversidade e Novidade. Essas métricas apresentaram valores idênticos para todas as abordagens comparadas, concluindo-se que o uso de STs não agregou valor ao algoritmo de recomendação.

No estudo de usuários, por outro lado, obteve-se bons resultados, com apenas uma das 48 recomendações feitas sendo descartada pelos usuários. Porém, ficou claro, a partir do *feedback* dado pelos usuários, que as publicações da rede de colaborações tiveram suas áreas classificadas incorretamente em grande parte dos casos. Mesmo com a classificação incorreta das publicações, ainda assim as recomendações foram bem recebidas em geral, mas acredita-se que isso se deve ao fato de os 50 pesquisadores contemplados pela abordagem atuarem na mesma instituição, e

portanto já conhecerem todos os pesquisadores recomendados.

Em trabalhos futuros pode ser considerado o uso de conjuntos de dados de pesquisadores de diversas instituições, possibilitando a geração de recomendações com alta serendipidade. Também pode-se considerar o uso de vários conjuntos de dados diferentes, para obtenção de mais dados estatísticos.

Quanto às áreas de pesquisa e à ontologia utilizada para classificar os artigos publicados, é sugerida a construção de uma nova ontologia, utilizando a versão 2012 do sistema de classificação da ACM (ou alguma versão mais recente no momento da publicação de novos trabalhos), e um novo dicionário de conceitos atualizado com dados de trabalhos publicados nos últimos anos.

A respeito do uso da Análise de Séries Temporais no contexto de recomendação de colaborações acadêmicas, ou no contexto de SRs em geral, acredita-se que há caminhos para se explorar em futuros trabalhos, embora os resultados iniciais deixem a desejar. É possível que se obtenha melhores resultados explorando mais profundamente os conceitos da STA e melhor adaptando seus modelos ao algoritmo de recomendação.

Por último, um experimento mais completo e de longo prazo é necessário para avaliar o impacto de um SR de colaborações acadêmicas no comportamento de grupos de pesquisadores em Ciência da Computação, de forma a se observar ao menos dois grupos distintos: um que irá receber as recomendações e outro que não irá receber. Este tipo de experimento permitirá que se obtenha dados mais precisos quanto à utilidade dos SRs e a satisfação de seus usuários.

REFERÊNCIAS

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **Knowledge and Data Engineering, IEEE Transactions on**, [S. l.]: IEEE, v. 17, n. 6, p. 734–749, 2005.
- BASILE, P. et al. Modeling short-term preferences in time-aware recommender systems. **Proceedings of DeCAT**, Dublin, Ireland: [s.n.], p. 44–54, 2015. Disponível em: <<http://ceur-ws.org/Vol-1388/DeCat2015-paper5.pdf>>. Acesso em: 5 abr. 2016.
- BOBADILLA, J. et al. Recommender systems survey. **Knowledge-Based Systems**, [S.l.]: Elsevier, v. 46, p. 109–132, 2013.
- BOSTANDJIEV, S.; O'DONOVAN, J.; HÖLLERER, T. Tasteweights: a visual interactive hybrid recommender system. **Proceedings of the sixth ACM conference on Recommender systems**, [S.l.]: ACM, p. 35–42, 2012.
- BRANDÃO, M. A.; MORO, M. M. Recomendação de colaboração em redes sociais acadêmicas baseada na afiliação dos pesquisadores. **Simpósio Brasileiro de Bancos de Dados**, São Paulo: [s.n.], 2012.
- BRANDÃO, M. A.; MORO, M. M.; ALMEIDA, J. M. Análise de fatores impactantes na recomendação de colaborações acadêmicas utilizando projeto fatorial. **Simpósio Brasileiro de Bancos de Dados**, Recife: [s.n.], 2013.
- BRANDÃO, M. A. et al. Using link semantics to recommend collaborations in academic social networks. **Proceedings of the 22nd international conference on World Wide Web companion**, [S.l.]: International World Wide Web Conferences Steering Committee, p. 833–840, 2013.
- BROCKWELL, P. J.; DAVIS, R. A. **Introduction to time series and forecasting**. [S.l.]: Springer Science & Business Media, 2006.
- BURKE, R. Hybrid recommender systems: Survey and experiments. **User modeling and user-adapted interaction**, [S.l.]: Springer, v. 12, n. 4, p. 331–370, 2002.
- BURKE, R. Hybrid web recommender systems. **The Adaptive Web**, [S.l.]: Springer, p. 377–408, 2007.
- CHAIWANAROM, P.; LURSINSAP, C. Collaborator recommendation in interdisciplinary computer science using degrees of collaborative forces, temporal evolution of research interest, and comparative seniority status. **Knowledge-Based Systems**, [S.l.]: Elsevier, v. 75, p. 161–172, 2015.
- CHATFIELD, C. **The analysis of time series: an introduction**. [S.l.]: CRC press, 2013.
- CHEN, C. et al. Terec: A temporal recommender system over tweet stream. **Proceedings of the VLDB Endowment**, [S.l.]: VLDB Endowment, v. 6, n. 12, p. 1254–1257, 2013.
- COGHLAN, A. **A Little Book of R For Time Series**. Release 0.2. [S.l.: s.n.], 2014. Disponível em: <<https://media.readthedocs.org/pdf/a-little-book-of-r-for-time-series/latest/a-little-book-of-r-for-time-series.pdf>>. Acesso em: 5 apr. 2016.

FOUSS, F.; SAERENS, M. Evaluating performance of recommender systems: An experimental comparison. **Web Intelligence and Intelligent Agent Technology**, [S.l.]: IEEE, v. 1, p. 735–738, 2008.

GAO, H. et al. Exploring temporal effects for location recommendation on location-based social networks. **Proceedings of the 7th ACM conference on Recommender systems**, [S.l.]: ACM, p. 93–100, 2013.

GE, M.; DELGADO-BATTENFELD, C.; JANNACH, D. Beyond accuracy: evaluating recommender systems by coverage and serendipity. **Proceedings of the fourth ACM conference on Recommender systems**, [S.l.]: ACM, p. 257–260, 2010.

HANANI, U.; SHAPIRA, B.; SHOVAL, P. Information filtering: Overview of issues, research and systems. **User Modeling and User-Adapted Interaction**, [S.l.]: Kluwer Academic Publishers, v. 11, n. 3, p. 203–259, 2001.

HERLOCKER, J. L. et al. Evaluating collaborative filtering recommender systems. **ACM Transactions on Information Systems (TOIS)**, [S.l.]: ACM, v. 22, n. 1, p. 5–53, 2004.

JANNACH, D. et al. **Recommender Systems: An Introduction**. [S.l.]: Cambridge University Press, 2010.

KNOKE, D.; YANG, S. **Social Network Analysis**. [S.l.]: Sage, 2008.

KONG, X. et al. Exploiting publication contents and collaboration networks for collaborator recommendation. **PLOS ONE**, [S.l.]: Public Library of Science, v. 11, n. 2, p. e0148492, 2016.

LATHIA, N. et al. Temporal diversity in recommender systems. **Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval**, [S.l.]: ACM, p. 210–217, 2010.

LEE, T. Q.; PARK, Y.; PARK, Y.-T. A time-based approach to effective recommender systems using implicit feedback. **Expert systems with applications**, [S.l.]: Elsevier, v. 34, n. 4, p. 3055–3062, 2008.

LI, J. et al. Acreec: a co-authorship based random walk model for academic collaboration recommendation. **Proceedings of the companion publication of the 23rd international conference on World wide web companion**, [S.l.]: International World Wide Web Conferences Steering Committee, p. 1209–1214, 2014.

LIKA, B.; KOLOMVATSOS, K.; HADJIEFTHYMIADES, S. Facing the cold start problem in recommender systems. **Expert Systems with Applications**, [S.l.]: Elsevier, v. 41, n. 4, p. 2065–2073, 2014.

LOH, S. et al. Constructing domain ontologies for indexing texts and creating users' profiles. **Work. on Ontologies and Metamodeling in Software and Data Engineering, Brazilian Symp. on Databases**, Florianópolis: [s.n.], n. 2003, p. 72–82, 2006.

LOPES, G. R. **Avaliação e Recomendação de Colaborações em Redes Sociais Acadêmicas**, 2012. 129 f. Tese (Doutorado em Ciência da Computação) – Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 2012.

LOPES, G. R. et al. Collaboration recommendation on academic social networks. **Advances in Conceptual Modeling – Applications and Challenges**, [S.l.]: Springer, p. 190–199, 2010.

LOPS, P.; GEMMIS, M. D.; SEMERARO, G. Content-based recommender systems: State of the art and trends. **Recommender systems handbook**, [S.l.]: Springer, p. 73–105, 2011.

MANNING, C. D. et al. **Introduction to information retrieval**. [S.l.]: Cambridge university press Cambridge, 2008.

MCNEE, S. M.; RIEDL, J.; KONSTAN, J. A. Being accurate is not enough: how accuracy metrics have hurt recommender systems. **CHI'06 extended abstracts on Human factors in computing systems**, [S.l.]: ACM, p. 1097–1101, 2006.

NAVARRO, G. A guided tour to approximate string matching. **ACM computing surveys (CSUR)**, [S.l.]: ACM, v. 33, n. 1, p. 31–88, 2001.

NEWMAN, M. E. Coauthorship networks and patterns of scientific collaboration. **Proceedings of the national academy of sciences**, [S.l.]: National Acad Sciences, v. 101, n. suppl 1, p. 5200–5205, 2004.

PAZZANI, M. J.; BILLSUS, D. Content-based recommendation systems. **The Adaptive Web**, [S.l.]: Springer, p. 325–341, 2007.

RICCI, F.; ROKACH, L.; SHAPIRA, B. **Introduction to recommender systems handbook**. [S.l.]: Springer, 2011.

RISTAD, E. S.; YIANILOS, P. N. Learning string-edit distance. **Pattern Analysis and Machine Intelligence, IEEE Transactions on**, [S.l.]: IEEE, v. 20, n. 5, p. 522–532, 1998.

SCHAFER, J. B. et al. Collaborative filtering recommender systems. **The Adaptive Web**, [S.l.]: Springer, p. 291–324, 2007.

SCHAFER, J. B.; KONSTAN, J.; RIEDL, J. Recommender systems in e-commerce. **Proceedings of the 1st ACM conference on Electronic commerce**, [S.l.]: ACM, p. 158–166, 1999.

SHANI, G.; BRAFMAN, R. I.; HECKERMAN, D. An mdp-based recommender system. **Proceedings of the Eighteenth conference on Uncertainty in artificial intelligence**, [S.l.]: Morgan Kaufmann Publishers Inc., p. 453–460, 2002.

SHANI, G.; GUNAWARDANA, A. Evaluating recommendation systems. **Recommender Systems Handbook**, [S.l.]: Springer, p. 257–297, 2011.

SHUMWAY, R. H.; STOFFER, D. S. **Time series analysis and its applications**. [S.l.]: Springer Science & Business Media, 2013.

ULLAH, F. et al. Hybrid recommender system with temporal information. **Information Networking (ICOIN), 2012 International Conference on**, [S.l.]: IEEE, p. 421–425, 2012.

WASSERMAN, S.; FAUST, K. **Social network analysis: Methods and applications**. [S.l.]: Cambridge university press, 1994.

XIA, F. et al. Mvwalker: Random walk-based most valuable collaborators recommendation exploiting academic factors. **Emerging Topics in Computing, IEEE Transactions on**, [S.l.]: IEEE, v. 2, n. 3, p. 364–375, 2014.

XU, Y. et al. Combining social network and semantic concept analysis for personalized academic researcher recommendation. **Decision Support Systems**, [S.l.]: Elsevier, v. 54, n. 1, p. 564–573, 2012.

ZHENG, N.; LI, Q. A recommender system based on tag and time information for social tagging systems. **Expert Systems with Applications**, [S.l.]: Elsevier, v. 38, n. 4, p. 4575–4587, 2011.

ZIEGLER, C.-N. et al. Improving recommendation lists through topic diversification. **Proceedings of the 14th international conference on World Wide Web**, [S.l.]: ACM, p. 22–32, 2005.