



XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017

GT-8 – Informação e Tecnologia

MODELAGEM CESM PARA UM SISTEMA DE RECOMENDAÇÕES: O CASO DE UMA LIVRARIA VIRTUAL

Cristina Elisabeth Ricken (Universidade Federal de Santa Catarina - UFSC)

Vinícius Medina Kern (Universidade Federal de Santa Catarina - UFSC)

CESM MODELING FOR A RECOMMENDER SYSTEM: A VIRTUAL BOOKSTORE CASE

Modalidade da Apresentação: Comunicação Oral

Resumo: Sistemas de recomendação têm, como objetivo, produzir recomendações personalizadas e capazes de guiar o usuário na escolha de um produto, serviço ou conteúdo em um grande espaço de opções possíveis. Por meio de filtragens por conteúdo e de filtragens colaborativas, um sistema de recomendações pode selecionar opções mais adequadas ao usuário, aliviando a sobrecarga de informações no contexto de livrarias virtuais. Neste sentido, o presente trabalho apresenta a modelagem e análise de um sistema de recomendações para uma livraria virtual vista como um sistema sociotecnológico. Esses sistemas envolvem agentes humanos e artificiais, suas interligações e dinâmica de interação, a sujeição às mudanças em relação às demandas ambientais e a emergência de comportamentos. Neste contexto, a abordagem metodológica é baseada na visão sistêmica de Mario Bunge, por meio da aplicação da modelagem CESM, segundo a qual um sistema pode ser representado conforme seus componentes, ambiente, estrutura e mecanismos. Os mecanismos envolvidos no sistema de recomendações são esboçados por meio de diagramas de causa-efeito. A modelagem CESM oferece uma visão integrada das interligações e interações entre os componentes de um sistema em seu contexto e a identificação dos processos envolvidos. Como o estudo de modelagem precede sua implementação computacional, podem ser realizados aprimoramentos e correções de aspectos não-funcionais relativos aos mecanismos antes de sua implantação no contexto de uma livraria virtual.

Palavras-Chave: Sistemas de recomendação; livraria virtual; filtragem híbrida; sistemas sociotecnológicos; modelo de sistema CESM.

Abstract: Recommender systems aim to create personalized recommendations for guiding the user in choosing the most useful product, service, or content in a large space of possible options. Through content filtering and collaborative filtering, a recommender system can select appropriate options for the user, alleviating the information overload in the context of virtual bookstores. In this sense, this work's aim is to present the modeling and analysis of a recommender system for a virtual bookstore seen as a sociotechnological system. A sociotechnological system involves human agents

and artificial agents, the interconnections between them, their interaction dynamics, subjection to changes in relation to environmental demands and the emergence of behaviors. In this context, the methodological approach is based on the systemic view of Mario Bunge, through the application of the CESM model, according to which a system can be represented by its components, environment, structure, and mechanism. Mechanisms involved in the recommender system are outlined as cause-effect diagrams. The CESM modeling offers an integrated view of the interconnections and interactions between the components of the system and their context, and the identification of the involved processes. The modeling study precedes its computational implementation and actions for improvements and corrections of non-functional aspects related to the mechanisms can be designed before its deployment in the context of a virtual bookstore.

Key-words: Recommender systems; virtual bookstore; hybrid filtering; sociotechnological systems; CESM system model.

1 INTRODUÇÃO

Sistemas de recomendação vêm sendo amplamente utilizados na Web, no comércio eletrônico e em plataformas de interação social para prover sugestões de itens ou de conteúdo com opções filtradas e personalizadas para cada usuário. Além de itens como livros, artigos científicos, vídeos e opções de entretenimento, sistemas de recomendação vêm trabalhando com o processamento e direcionamento de conteúdo indicado ou gerado pelos próprios usuários (LI et al., 2010; BELLOGÍN; CANTADOR; CASTELLS, 2013). Dessa forma, ampliando o conceito de sistemas de recomendações para a dimensão social, ambientes computacionais que incorporam opções de compartilhamento e colaboração podem processar os dados e as informações extraídas, com sua conversão em recomendações com maior chance de suscitar interesse e atender às necessidades dos usuários.

No presente trabalho, um sistema de recomendações para uma livraria virtual será modelado e analisado como um sistema sociotecnológico – baseado na colaboração dinâmica humano-máquina (ECCLES e GROTH, 2006; KERN, 2011), segundo a visão sistêmica de Mario Bunge (BUNGE, 2003; 2004; 2006) e seu modelo CESM (*Composition, Environment, Structure, Mechanism*) (BUNGE, 2003). Essa modelagem precede o desenvolvimento do sistema computacional propriamente dito e abrange os possíveis comportamentos do sistema diante das ações e interações dos elementos envolvidos.

Um sistema de recomendação pode ser visto como um sistema sociotecnológico, na medida em que agentes humanos (usuários) e agentes tecnológicos (agentes artificiais, em especial software) coordenam suas atividades nos processos de recomendações sobre livros. Além de buscas, navegação e aquisições, os usuários mobilizam o ambiente do sistema de recomendações ao avaliar os livros, publicar resenhas e comentários, assim como ao atribuir utilidade percebida às contribuições do conteúdo gerado e publicado por outros usuários. Os conjuntos de dados obtidos de forma implícita (como comportamento de navegação e de aquisição no sítio da livraria) e de forma explícita (como atribuição de *ratings* – escores de avaliação e de utilidade por parte do usuário) são utilizados por algoritmos implementados para o cômputo de recomendações pelo sistema computacional.

Os componentes tecnológicos nos sistemas sociotecnológicos são planejados, teleológicos. Porém, as interações entre pessoas e agentes tecnológicos, como condição para o funcionamento do sistema como um todo, podem dar emergência a novas condições de

desempenho não previstas no desenho inicial do sistema tecnológico. Como consequência destas interações, o sistema tende à auto-organização em resposta às mudanças ambientais que o afetam, em um processo do qual podem emergir relações não-lineares e complexas.

Sistemas sociotecnológicos, como um sistema de recomendações para livraria virtual, não são modeláveis ou analisáveis por métodos usuais da tecnologia, que pressupõem funcionamento mecanístico. A análise do sistema segundo o modelo CESM de Bunge (2003) permite explicar a emergência em um sistema sociotecnológico por meio de diagramas de Boudon-Coleman (BUNGE, 2004), que permitem conjecturar e investigar as causas da emergência, manutenção ou decaimento de propriedades. A análise e investigação de mecanismos e processos do sistema será feita utilizando conceitos de dinâmica de sistemas e diagramas de causa-efeito (*causal loop diagrams – CLD*) (STERMAN, 2000).

O artigo prossegue com a apresentação dos conceitos essenciais sobre um sistema de recomendação incorporável a uma livraria virtual (Seção 2); com a modelagem do sistema de recomendação para livraria virtual, incluindo a descrição de seus componentes, itens ambientais e estrutura e o delineamento da investigação de seus mecanismos (Seção 3) e finaliza com as Considerações Finais do trabalho (Seção 4).

2 SISTEMAS DE RECOMENDAÇÃO

Um sistema de recomendação é qualquer sistema que produz, como saída, recomendações individualizadas ou que possui o efeito de guiar o usuário de forma personalizada no processo de escolha de algum item (produto, serviço ou conteúdo) no espaço de itens possíveis (BURKE, 2002). Sistemas de recomendação podem aliviar tanto a sobrecarga quanto a falta de informação, apresentando ao usuário um conjunto de itens filtrados para este. Os itens recomendados podem ser computados a partir de requisições e das preferências explicitadas ou deduzidas do usuário a partir de características de itens ou de assuntos buscados e com base em ações e opiniões de outros usuários.

O sistema de recomendação é capaz de prover sugestões ou recomendações a um usuário em seu contexto de demanda. Conforme Burke (2002), Adomavicius e Tuzhilin (2005), as recomendações realizadas podem ser baseadas em: dados demográficos, conhecimento, filtragem por conteúdo (FCont) e filtragem colaborativa (FColab). A filtragem híbrida pode combinar FCont e FColab, além de filtragens por meio de dados demográficos, utilidade e conhecimento.

2.1 Filtragem por conteúdo

Métodos de recomendação baseados em FCont utilizam informações sobre os atributos descritores de um item e dos *ratings* (escores) atribuídos a este por um usuário. Algoritmos efetuam a análise automatizada de conteúdo de itens e agentes de software realizam a correlação entre o conteúdo dos itens já avaliados (para construir um perfil do usuário) e o conteúdo de um item novo. Esta técnica combina os escores atribuídos com um perfil de interesses do usuário baseado, por exemplo, nos atributos dos itens avaliados. A partir disso, o sistema de recomendação pode encontrar itens com os atributos aferidos e recomendar os que possuem maior similaridade com aqueles preferidos pelo usuário no passado. As recomendações baseadas exclusivamente em conteúdo baseiam-se nas informações individuais de cada usuário sem considerar o comportamento observado e as avaliações realizadas por outros usuários. Para realizar boas predições de preferência para o usuário, FCont depende de histórico significativo de buscas e aquisições e, portanto, as estimativas para um usuário novo, sem histórico e perfil, tendem a ser menos acertadas.

2.2 Filtragem colaborativa

Sistemas de recomendação que utilizam FColab baseiam o cômputo das recomendações para um determinado usuário nas preferências demonstradas por outros usuários. Os itens são processados e selecionados com base na concordância passada de ações e avaliações de usuários cujo perfil é computado como similar ao do usuário interessado.

Os dados utilizados na FColab são obtidos de forma explícita e implícita. Dados são coletados de forma explícita, através dos *ratings* atribuídos aos itens (livros) avaliados por outros usuários para estimar preferências para o usuário em questão. Dados podem ser coletados de forma implícita, através do histórico de navegação e de consumo de usuários, conduzindo às recomendações na forma de: “usuários que buscaram/compraram x, também buscaram/adquiriram livros y, z, w, ...”. Ainda que dispensem a análise de conteúdo de um item, as recomendações a serem realizadas dependem de um conjunto suficiente de *ratings* de outros usuários para o cômputo da predição de utilidade de um item novo para um determinado usuário.

2.3 Filtragens Híbridas

FColab e FCont são abordagens que vêm sendo intensamente estudadas em sistema de recomendação (BURKE, 2002; ADOMAVICIUS; TUZHILIN, 2005; BOBADILLA, 2013). No entanto, existem algumas limitações associadas a estas que interferem na qualidade das recomendações realizadas e/ou na eficiência do sistema, como o problema de esparsidade (SARWAR, 2000), em que uma matriz (p. ex. matriz usuário-item) se torna muito esparsa para o cálculo de similaridades, e o problema de escalabilidade, quando existe grande volume de usuários e de itens, tornando o processamento e a análise muito onerosos. A Filtragem híbrida (BURKE, 2002; ADOMAVICIUS; TUZHILIN, 2005) pode auxiliar na supressão de algumas das limitações associadas à adoção exclusiva de FColab ou de FCont ao combinar duas ou mais abordagens de filtragem.

3 MODELAGEM CESM PARA O SISTEMA DE RECOMENDAÇÃO

A pesquisa desenvolvida neste trabalho é caracterizada como exploratória, descritiva e qualitativa. Metodologicamente, é aplicado o modelo CESM (BUNGE, 2003) como uma modelagem qualitativa de um sistema sociotecnológico que precede a concepção e o desenvolvimento da plataforma computacional que o sustenta. Permite investigar, por meio de simulações, os possíveis comportamentos do sistema diante das ações e interações dos elementos envolvidos tendo, como base, conhecimento teórico e empírico dos aspectos contextuais em que está inserido.

Como objetivo, o sistema de recomendações integrado à livraria virtual visa proporcionar aumento das aquisições de livros por meio de recomendações adequadas que culminem com a conversão das buscas em aquisições efetivas. Além deste aspecto, o sistema de recomendação procura manter uma base de usuários regulares e que são vetores de recomendação da própria livraria. O sistema de recomendação modelado e analisado no presente trabalho consiste em uma concepção fictícia, porém inspirada em características observáveis de sistemas de recomendação de livrarias virtuais na Web.

3.1 Modelagem CESM

A modelagem sistêmica CESM (*Composition, Environment, Structure, Mechanism*), proposta por Bunge (2003; 2004; 2006), permite a representação dos elementos de um sistema, suas interações e a conjectura dos mecanismos que o regem. O CESM é definido segundo um modelo $\mu(\sigma)$ de um sistema σ , em função da Composição $C(\sigma)$, Ambiente $E(\sigma)$, Estrutura, $S(\sigma)$ e Mecanismo $M(\sigma)$, sendo expresso por: $\mu(\sigma) = \langle C(\sigma), E(\sigma), S(\sigma), M(\sigma) \rangle$

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

$C(\sigma)$ - Composição - Consiste na coleção de elementos (Componentes) que compõem o sistema, sendo que aqueles que estabelecem ligações com o ambiente em que este está inserido formam a fronteira do sistema.

$E(\sigma)$ - Ambiente - Diz respeito à coleção de elementos que não são parte do sistema, mas que atuam sobre Componentes do sistema e/ou recebem ações destes. Bunge ressalta que todos os sistemas, com exceção do universo, têm um ambiente.

$S(\sigma)$ - Estrutura - Abrange o conjunto de conexões entre componentes constituintes do sistema, assim como as ligações entre os componentes e os elementos do ambiente. As conexões são fundamentais para a auto-organização do sistema. A estrutura pode ser ainda definida em termos da Endoestrutura (coleção de ligações entre componentes) e da Exoestrutura (coleção de ligações dos componentes com itens do ambiente) do sistema.

$M(\sigma)$ - Mecanismo - Corresponde ao conjunto de processos que tornam o sistema o que este é. Mecanismos afetam o comportamento do sistema, provocando mudanças, seja na forma de emergência ou de decomposição de uma parte do sistema ou de algumas de suas propriedades.

A modelagem CESM para um sistema completo com todos os seus subsistemas e interações pode tornar-se bastante complexa. Assim, de acordo com Bunge (2003), para uma modelagem sistêmica prática, os modelos podem ser desenhados em um certo nível *a* da composição, nível *b* do ambiente, nível *c* da estrutura, e nível *d* do mecanismo. Deste modo, diferentes modelos de um mesmo sistema podem ser construídos, variando os níveis *abcd* da análise para os modelos de subsistemas deste, conforme o corte de nível definido.

Como um sistema sociotecnológico, o corte do sistema de recomendações da livraria virtual, em análise, é realizado no nível das relações entre os componentes: agentes humanos (usuários do sistema); o livro, que consiste no denominado artefato essencial (KERN, 2011) e os agentes de software que automatizam operações no processo de recomendação. Por meio do metamodelo CESM é facultada uma visão integrada das interligações e interações entre os componentes do sistema de recomendação em seu contexto de operação, assim como a identificação do modo de funcionamento dos processos envolvidos para sua correção e aprimoramento. Deste modo, é viabilizado o foco da análise neste sistema e nas inter-relações entre seus subsistemas componentes, sendo que as transações entre itens individuais dentro de um determinado subsistema componente não são detalhadas. Da mesma forma, ligações entre itens do ambiente (cada qual formando um outro sistema) são reconhecidas, mas não

analisadas, e é respeitada a devida proporção de influência entre componentes do sistema de recomendação e os itens do ambiente em que está inserido.

3.2 Modelagem CESM para o sistema de recomendação para livraria virtual

O sistema de recomendações tem, como ambiente: a Web; infra-estrutura TIC; a livraria virtual (como organização); os regulamentos de editoras; inovações tecnológicas; a cultura e o mercado da livraria virtual. Estes elementos ambientais mantêm ligações com alguns componentes do sistema no nível da exoestrutura. No nível atual da modelagem do sistema de recomendação, os usuários humanos são agregados e o conjunto é tratado como um subsistema componente denominado de i) agentes humanos. Ou seja, em lugar de examinar comportamentos no nível individual, será examinado o efeito conjunto das interações entre agentes humanos. Analogamente, os agentes artificiais formam um subsistema multi-agentes, sendo este considerado um componente do sistema denominado de ii) agentes de software. Neste sentido, é reconhecido o conjunto de ligações entre estes, mas não são examinadas as transações entre agentes de software, mas o efeito das ações destes no processo de recomendação. Os livros, por sua vez, formam o subsistema componente iii) artefato essencial. A emergência de comportamentos decorre das interligações entre estes componentes e de sua dinâmica de interação no processo de recomendação.

O quadro-resumo, descrito na Fig. 1, sintetiza os elementos e aspectos do Sistema em análise, segundo um modelo CESM.

Quadro 1: Síntese do CESM aplicado ao sistema de recomendação para livraria virtual

CESM	Descrição
Sistema	Sistema Sociotecnológico de Recomendação em Livraria Virtual
Composição	Livros (artefato essencial); agentes humanos; agentes de software,
Ambiente	Web; infra-estrutura de tecnologia da informação e comunicação; livraria virtual (como organização); regulamentos de editoras; cultura; inovações tecnológicas; mercado da livraria
Estrutura	Navegação na busca por livros no sítio da livraria; recomendações; <i>ratings</i> ; mensagens; comentários; efetuação de compra; coleta de dados associados ao livro - <i>ratings</i> ; coleta de dados para construção de perfis de usuários; coleta de dados para análise de conteúdo, mineração de dados e de texto; comunicação e coordenação entre agentes nas operações de formação de perfis (modelos) de usuários, nas filtragens e na formação de <i>ranks</i> de itens; ranqueamento de itens; similaridade entre livros; <i>links</i> do livro para a web; <i>links</i> da Web para o livro.
Mecanismos	Filtragens híbridas de informação; colaboração humana por meio de <i>ratings</i> para filtragem colaborativa de informação; colaboração humana por meio de avaliações e revisões de forma textual.

Fonte: Os autores

O sistema de recomendações da livraria virtual modelada combina filtragens baseadas em FCont e FColab, caracterizando uma Filtragem híbrida. Além destas filtragens, as contribuições humanas na forma de avaliações e revisões de forma textual também são consideradas no sistema e também tendem a afetar a decisão de escolha e de aquisição por parte de um indivíduo.

3.3 Composição

No nível de corte para modelagem e análise do sistema são considerados como componentes essenciais, seguindo o CISM (BUNGE, 2003; KERN, 2011): os livros que compõem o catálogo da livraria e que constituem o i) artefato essencial que estabelece conexões entre agentes humanos e agentes de software; os ii) agentes humanos (usuários da livraria virtual); e os iii) agentes de software que automatizam coleta de dados, processam os algoritmos para filtragens e realizam operações envolvidas na produção de recomendações.

3.3.1 C1 Livro como Artefato Essencial

Cada livro (título único) está associado com: estatísticas relativas às buscas; número de unidades vendidos; *ratings* atribuídos por usuários; conjunto de atributos descritores; relação de similaridade com outros livros, conforme conteúdo; dados de preço, disponibilidade; relação com outros livros e com usuários por meio de busca, navegação e aquisição.

3.3.2 C2 Agentes Humanos

O conjunto de agentes humanos é formado pelos usuários e clientes que: realizam buscas, navegação e aquisições (que produzem dados implícitos); atribuem *ratings* aos livros (avaliações sob a forma de pontos em uma escala crescente de valor para inteiros de 1 a 5) que indicam grau de apreciação e são utilizados para geração de recomendações (dados obtidos de forma explícita); contribuem por meio de resenhas e comentários sobre livros (avaliações em forma textual), associados aos *ratings* atribuídos aos livros; respondem de forma textual a comentários feitos sobre livros; atribuem utilidade aos comentários feitos (contabilizando o total de “este comentário foi útil” emitidos)

3.3.3 C3 Agentes de Software

Agentes de software consistem em entidades computacionais especializadas em determinadas funções e que são integradas em um sistema para executar tarefas que auxiliam agentes humanos em uma determinada atividade (JENNINGS; SYCARA; WOOLDRIDGE, 1998). No nível atual do sistema de recomendação modelado e analisado, os diversos agentes

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

artificiais envolvidos são agregados para formar um subsistema considerado como um componente denominado “agentes de software”. Assim, são descritas funções que agentes artificiais têm no sistema e sua relação com outros componentes, mas sem especificar precisamente cada agente e sem examinar as transações entre estes.

No sistema de recomendação, agentes de software realizam: operações gerais de coleta de dados; tarefas de extração e agregação de informações de dados; modelagem de perfis de usuários; extração de informação e mineração de conteúdo; operações de filtragens envolvidas na produção de recomendações; o ranqueamento de livros conforme o seu grau de utilidade computada. Para realizar estas tarefas, são utilizados dados associados aos usuários (*ratings* atribuídos a cada livro); dados relativos à navegação e transações, aquisições; dados sobre os livros (atributos dos livros, extrações e mineração dos textos, metadados).

Agentes envolvidos nas filtragens. Nos processos de filtragens (colaborativa e por conteúdo) os agentes utilizam dados que permitem construir um perfil para o usuário, aplicado para estimar suas preferências em relação aos livros. Podem ser utilizados dados obtidos de modo explícito, através de *ratings* para os livros, assim como dados obtidos através do comportamento de navegação e de transações ao longo do tempo. Para construção de perfis de usuários, na FCont e na FColab podem ser implementadas técnicas derivadas de recuperação de informações, mineração de texto, mineração de dados e de aprendizado de máquina.

Filterbots. São agentes de software que provêm avaliações automáticas para os itens (BURKE, 2002) e que podem realizar análise de conteúdo destes. Estes agentes artificiais atuam como participantes adicionais em uma comunidade de FColab, provendo mais avaliações nas quais as recomendações podem se basear. Assim, usuários cujos *ratings* concordam com alguns dos *ratings* dos *Filterbots* têm maiores chances de receber melhores recomendações. Agentes capazes de aprendizado podem ser personalizados para um usuário individual, atualizando seu perfil ao longo da interação deste com o sistema.

Agente recomendador. As recomendações efetuadas por este agente assumem a forma de listas de livros que são frequentemente adquiridos por outros usuários que também buscaram por um dado livro. Também são apresentados livros com características de similaridade de itens preferidos pelo usuário. Estas recomendações se baseiam, tanto na FCont, quanto na comparação de perfis de usuários por meio de FColab. Na página de resultados também são apresentados os *ratings* atribuídos, resenhas e comentários realizados

por outros usuários sobre o livro em questão. Estas contribuições de usuários em forma textual possuem teor de validade que é julgado pelo próprio usuário que busca informações sobre o livro em questão.

3.4 Ambiente

O ambiente para o sistema de recomendação de livros é formado pela livraria virtual, como organização e supersistema; regras e regulamentos de editoras; pelo mercado da livraria; pela infraestrutura de Tecnologia da Informação e da Comunicação que viabiliza o sistema; pelas inovações tecnológicas; pela Web e pela Cultura do ambiente.

A1. Livraria virtual. Sedia o sistema de recomendações, as bases de livros e de usuários; avalia o mercado, toma decisões táticas e estratégicas quanto ao negócio; gere contratos e efetua negociações com editoras de livros; examina e incorpora inovações tecnológicas que afetam as buscas por livros; elabora e gere políticas de privacidade e de segurança; elabora promoções dirigidas (preços, vendas casadas, surpresas, etc ...).

A2. Regras e regulamentos de editoras. Têm pouca influência direta no funcionamento do sistema de recomendação propriamente dito, mas têm impacto indireto, através da livraria virtual, sobre a busca e aquisição de livros. Afetam o negócio da livraria virtual e suas decisões de compra e promoções de livros, assim como o grau de exposição sobre seu conteúdo. Os acordos entre editoras e a livraria virtual determinam a exposição ou não de parte do conteúdo do livro no site, fornecendo ao usuário mais informações sobre o livro buscado, o que tende a afetar a decisão de compra.

A3. Mercado da livraria virtual. Abrange parcerias e concorrência ao negócio da livraria virtual e o mercado produtor e consumidor de livros. Afeta o acesso/tráfego, oferta/demanda, preços/promoções e inclusão/exclusão de obras na base de dados da livraria.

A4. A infraestrutura TIC para processamento, armazenamento e transferência dos dados. Fornece a estrutura de uso do sistema de recomendação, sua ligação com as bases de dados de livros, usuários; viabiliza acesso e transações seguras para o usuário e o uso deste meio por agentes de software e livros.

A5. Web. Apresentação nos resultados de busca por meio de engenhos de buscas; ligações com outros sites, blogs, etc ... Difusão de informações e tendências relacionadas às obras/livros, como notícias, filmes, eventos em geral.

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

A6. Cultura. A cultura na qual usuários estão imersos influencia as buscas por livros e tendências de interesse, afetando, também, decisões da livraria sobre os livros que formarão a base de dados. O desenvolvimento de uma cultura de colaboração e compartilhamento entre pessoas possivelmente afeta o comportamento de colaboração de agentes humanos, por meio de avaliações (revisões e atribuições de *ratings*) no sistema de recomendações.

A7. Inovações tecnológicas. Leitores de livros eletrônicos, por exemplo, suscitam iniciativas de promoções de vendas de livros associados a estes. Assume-se que o investimento do usuário em um determinado leitor eletrônico tenda a favorecer sua adesão às promoções da livraria virtual, a intensificar o tráfego em seu ambiente e o aumento de vendas dos livros, o que, por sua vez, alimenta as bases de dados sobre navegação e sobre consumo, propiciando a geração de recomendações para livros.

3.5 Estrutura

A estrutura é definida em termos de: endoestrutura (coleção de ligações entre componentes) e exoestrutura (coleção de ligações dos componentes com itens do ambiente). São componentes da endoestrutura: Os livros (artefatos essenciais), designados por (C1), os agentes humanos por (C2), agentes de software; agentes de software por (C3). A exoestrutura é formada pelas ligações entre os Componentes do sistema e os elementos do Ambiente, como: (A1) livraria virtual; (A2) regulamentos e regras de editoras; (A3) mercado da livraria virtual; (A4) infraestrutura TIC; (A5) Web; (A6) cultura; (A7) inovações tecnológicas.

As ligações entre os agentes humanos/usuários (C2) do sistema ocorrem através de comentários relativos às revisões, resenhas em forma textual e da atribuição de utilidade a estas. Usuários se ligam aos artefatos essenciais (C1) por meio das buscas, da conversão destas em compras, dos *ratings* associados aos comentários/resenhas sobre livros. Como artefato essencial no sistema, um livro é apresentado ao usuário por meio de atributos (título, autor, editora, data sumário, preço, etc ...), acompanhado de outros livros (C1) com os quais este tenha alguma relação de similaridade identificada pelo sistema de recomendação.

Os agentes de software (C3) se ligam aos usuários por meio da coleta de dados e de estatísticas utilizados para compor o perfil de cada qual (*ratings* associados aos livros, preferências, relações com outros usuários, análise de comportamento de navegação e de suas transações ao longo do tempo) e da apresentação das recomendações. Agentes de software (C3) se comunicam e cooperam entre si nas operações de construção de perfis de

**XVIII ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2017
23 a 27 de outubro de 2017 – Marília – SP**

usuários, de *data mining* e nas filtragens realização. Realizam a análise de conteúdo dos livros (C1), assim como a coleta de dados para extração e análise de estatísticas associadas a cada item, como *ratings*, e a apresentação dos itens computados como adequados ao usuário, na forma de *ranks*.

As ligações entre usuários (C2) e Livraria virtual (A1) ocorrem por meio de promoções dirigidas ao perfil de consumo de usuário; de políticas de privacidade e de segurança quanto aos seus dados. A1 se liga à C3 por meio de Agentes que realizam a extração processamento de dados das bases da livraria. A1 também toma decisões que afetam C1, quanto ao Catálogo de livros e a exposição ou não de conteúdo de forma gratuita afeta a busca/compra de livros.

As regras e regulamentos de editoras (A2) quanto à distribuição gratuita de conteúdo tem impacto sobre a busca e compra de livros. Os acordos entre editoras e a livraria virtual determinam a exposição ou não de parte do conteúdo do livro no site, fornecendo ao usuário mais informações sobre o livro buscado. O Mercado da livraria virtual (A3) afeta a oferta, demanda e preços e a inclusão/exclusão de obras na base de dados. A infraestrutura TIC (A4) viabiliza o acesso e a navegação do usuário do sistema, e o uso do meio por livros (C1) e agentes de software (C3).

A Web (A5) propicia a difusão de informações e tendências que afetam buscas pelos usuários e também apresenta obras nos resultados de engenhos de buscas, e com a qual livros mantêm ligações. A Cultura (A6) do ambiente no qual usuários estão imersos influencia o interesse das buscas e o comportamento de colaboração entre agentes humanos no sistema, influenciando, também, decisões sobre os livros que formarão a base de dados. Inovações tecnológicas (A7), como leitores de livros eletrônicos, afetam decisões sobre o aumento do catálogo de livros (C1) e promoções. O investimento do usuário (C2) em tecnologia compatível favorece sua adesão às promoções da livraria virtual. A7 também influencia os aprimoramentos sobre agentes de software (C3) incorporados ao sistema de recomendações.

3.6 Mecanismos

No contexto do CESM, mecanismos são processos que causam emergência, manutenção, decaimento ou destruição de propriedades do sistema, permitindo explicar o seu funcionamento (BUNGE, 2004). Geralmente, mecanismos precisam ser conjecturados, o que pode ser realizado por meio da identificação de funções do sistema para que se possa inferi-

los. Para o sistema modelado e analisado, o foco recai sobre o comportamento e o desempenho do sistema no processo de geração de recomendações.

A função essencial de um sistema de recomendação é gerar recomendações adequadas aos usuários, por meio da utilização de recursos (dados, algoritmos/métodos de filtragem, infraestrutura computacional) disponíveis. Recomendações são consideradas satisfatórias/interessantes, conforme o julgamento subjetivo do usuário. Já, a adequação das recomendações é alcançada dentro do limite de recursos de operação do sistema. Ou seja, métodos de filtragem pouco eficientes e pouco acurados, assim como dados de baixa qualidade, tendem a gerar recomendações inadequadas.

Uma vez que um usuário inicie suas buscas no sítio da livraria, o sistema deve realizar recomendações, selecionando livros que sejam potencialmente adequados a este, ao longo de um catálogo de livros. Os mecanismos que viabilizam a produção de resultados potencialmente adequados ao usuário são, essencialmente, as filtrações computacionais (FColab, FCont, a conjugação destas e de outras técnicas) e a colaboração, por meio de avaliações de usuários. Técnicas de filtrações são operacionalizadas por meio de métodos/algoritmos e utilizam dados implícitos e explícitos disponíveis. A colaboração dos usuários se dá de três maneiras: a) por meio da atribuição de *ratings* aos livros b) de forma textual, por meio de resenhas, de comentários e revisões de livros c) por meio da utilidade atribuída a estas ações de a) e b). A seguir são apresentadas as conjeturas dos mecanismos atuantes no nível de corte do sistema, por meio de Diagramas de Boudon-Coleman (BUNGE, 2003).

3.6.1 Filtrações Híbridas

O mecanismo relativo às Filtrações é essencial ao funcionamento do sistema de recomendação pois é através deste que são selecionados, a partir de um vasto catálogo, os livros estimados como adequados ao usuário. O termo Filtrações Híbridas designa estas filtrações que englobam: os métodos de FCont; de FColab; além da aplicação de *Filterbots* e de técnicas de redução de dimensionalidade e tratamento de esparsidade de matrizes, dentre outras, para realizar as predições que conduzem às recomendações.

Filtrações Híbridas procuram balancear qualidades de diversos métodos, compensando limitações que ocasionam recomendações pouco adequadas e baixa eficiência do sistema de recomendações. Como exemplo de limitação associada à FCont pode ser citado o excesso de

especialização em relação ao perfil do usuário e que impede a recomendação de livros que representem novidade para este. Ao utilizar perfis de outros usuários, a FColab pode compensar esta limitação de FCont, mas também apresenta suas próprias limitações, como o problema de escalabilidade (número muito grande de usuários e de itens, tornando o processo de análise e comparação de perfis computacionalmente oneroso) e o problema de esparsidade (caracterizado por um número insuficiente de avaliações, tornando a matriz usuário-item muito esparsa e dificultando a realização de recomendações adequadas).

Figura 1: Mecanismo Filtragens Híbridas.



Fonte: Os autores

Ao combinar diversas formas de filtragem, a aplicação de Filtragens híbridas complementa as qualidades de cada técnica e atenua suas limitações. A melhor utilização dos recursos do sistema auxilia na predição de utilidade de livros para o usuário, aumentando o potencial preditivo do sistema de recomendação. Desta forma, Filtragens Híbridas contribuem para a produção de recomendações adequadas ao usuário, conforme os dados de que dispõe o sistema.

3.6.2 Contribuição humana por meio de ratings para Filtragem colaborativa

A filtragem colaborativa que utiliza dados explícitos baseia as recomendações a partir da predição de utilidade de um determinado item para o usuário em questão, conforme os *ratings* previamente atribuídos aos livros por outros usuários. Assim, a utilidade $f(u, l)$ de um livro l para um usuário u é estimada com base nas utilidades $f(u_i, l)$ atribuídas ao livro l pelos usuários u_i do conjunto U de usuários similares ao usuário u . Uma matriz usuário-item é construída, a partir da qual são selecionados itens com base em algum critério de similaridade de perfil entre outros usuários e o perfil do usuário interessado. No entanto, o número de *ratings* disponíveis, geralmente, é muito menor que o número necessário para realizar as predições, conduzindo ao problema de esparsidade.

Figura 2: Mecanismo Contribuição humana por meio de *ratings* para Filtragem colaborativa.



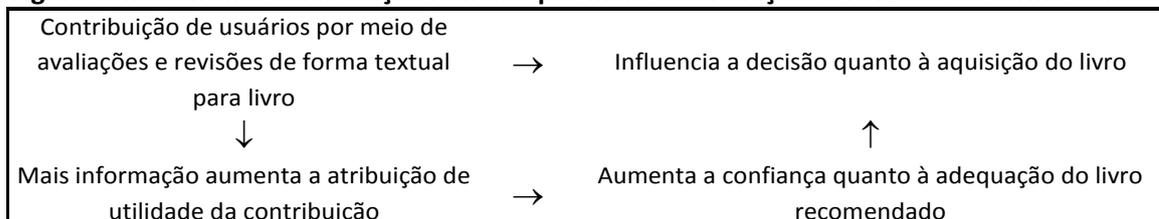
Fonte: Os autores

A contribuição de usuários u_i por meio da atribuição de *ratings* aos livros auxilia na redução de esparsidade da matriz usuário-item em FColab. Com maior número de *ratings* que se sobrepõem para livros avaliados, é possível melhorar o cômputo da similaridade entre perfis de usuários u_i que avaliaram os livros e o perfil do usuário u . Esta forma de colaboração humana incorporada à filtragem colaborativa por meio de técnicas computacionais contribui para a produção de recomendações potencialmente mais adequadas para o usuário u . O comportamento de abstenção nas avaliações, oposto ao da colaboração, pode aumentar a esparsidade da matriz usuário-item, a ponto de tornar inviável o cômputo de similaridade entre usuários e a produção de recomendações adequadas para estes. Ressalta-se que esta forma de colaboração humana depende da confiabilidade das avaliações por meio de *ratings* para geração de estimativas de utilidade para o usuário u .

3.6.3 Contribuição humana por meio avaliações e revisões de forma textual

Ratings (de 1 a 5) emitidos por usuários para livros podem ser acompanhados de contribuições (resenhas, comentários, avaliações) em forma textual que elucidam os *ratings* atribuídos. Assume-se que o número de usuários que avaliaram um livro e os *ratings* atribuídos e às contribuições têm influência na percepção de sua utilidade por parte de usuários que buscam pelo livro em questão.

Figura 3: Mecanismo Colaboração humana por meio de avaliações e revisões de forma textual.



Fonte: Os autores

A contribuição de usuários por meio de avaliações (*ratings*) e de revisões, resenhas, comentários em forma textual para um livro fornece mais informação que justifique os *ratings* atribuídos. Mais informações, neste sentido, aumentam a atribuição de utilidade à contribuição e aumentam a confiança quanto ao grau de adequação do livro recomendado

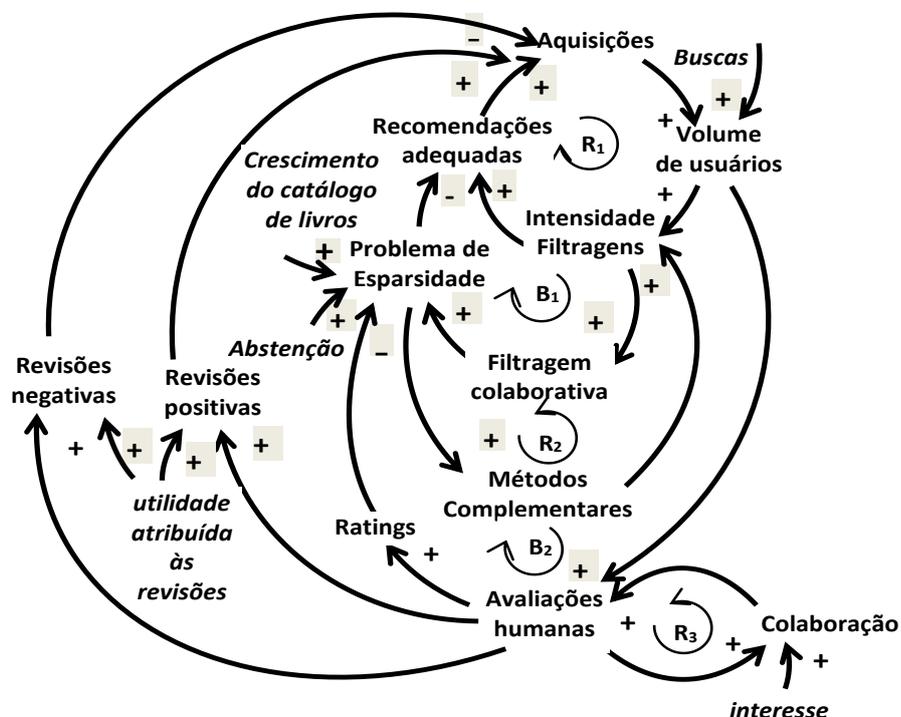
para o usuário. A diminuição de incerteza por meio de informações, recomendações textuais de outros usuários e maior confiança na opinião formada pelo usuário buscador contribuem para a decisão sobre aquisição (ou não) do livro.

3.7 Delineamento dos Mecanismos através de Diagramas de Causa-Efeito

O estudo investigativo de mecanismos se concentrará no processo de emissão de recomendações adequadas por meio de filtros e de avaliações humanas e sua conversão em aquisições, com foco no problema de esparsidade que decorre da filtragem colaborativa. Utiliza conceitos da dinâmica de sistemas com diagramas de causa-efeito (STERMAN, 2000).

O aumento do número de usuários e de itens conduz ao problema da esparsidade na filtragem colaborativa (SARWAR et al., 2000). No sistema de recomendações essa esparsidade é intrínseca ao fato de que, ainda que haja um grande número de usuários, é baixo o percentual de usuários que avalia os livros, devido à enorme quantidade de livros que compõem o catálogo e ao comportamento humano de se abster na avaliação de livros. A Figura 4 ilustra a modelagem dinâmica por meio de *loops* de causa-efeito para este cenário.

Figura 4. Modelagem dinâmica por meio de *loops* de causa-efeito



Fonte: Os autores

Loop de reforço R1 (Figura 4). O aumento do **Volume de usuários** intensifica as Filtragens para a provisão de **Recomendações Adequadas** que contribuem para o aumento

das **Aquisições**. O aumento de **Aquisições** e as *Buscas* realizadas fazem crescer o **Volume de usuários** e de seus dados (relativos a itens buscados, navegação e aquisições), aumentando a **Intensidade de Filtragens** que possam selecionar itens adequados aos usuários.

Loop de balanço B1. Com o aumento da **Intensidade de Filtragens**, em decorrência do crescimento do **Volume de usuários**, também aumenta a **Filtragem Colaborativa** (FColab). Na FColab há a necessidade de *ratings* suficientes na matriz usuário-item para comparação de similaridade de perfis de usuários. No entanto, com *Crescimento do catálogo de livros*, associado à *Abstenção* por parte de usuários que não emitem *ratings*, aumenta o **Problema da Esparsidade** da matriz usuário-item. O **Problema de Esparsidade** reduz a geração de recomendações adequadas com base nos recursos disponíveis, demandando correções.

Loop de reforço R2. Com a Esparsidade da matriz usuário-item, aumenta a aplicação de **Métodos de Filtragem Complementares** para melhorar a utilidade da FColab na produção de recomendações (Filtragens são complementadas por meio de FCont, adoção de Filterbots, tratamento da esparsidade da matriz usuário-item). O esforço em **Métodos complementares**, por sua vez, também aumenta a **Intensidade das Filtragens** em geral.

Loop de balanço B2. As **Avaliações humanas** assumem a forma de *ratings* e de revisões textuais realizadas por usuários. O aumento na quantidade de **Avaliações humanas** se dá em decorrência do aumento no **Volume de Usuários** e da **Colaboração (Loop R3)** destes, ativada pelo *interesse* em colaborar. O aumento de **Avaliações Humanas** na forma de *ratings* contribui para a diminuição da esparsidade da matriz usuário-item.

O **Loop de reforço R3** ilustra como o comportamento de **Colaboração** aumenta a quantidade de **Avaliações Humanas**, o que realimenta a **Colaboração** (por exemplo, avaliações humanas textuais tendem a suscitar discussões que mobilizam contribuições). O *interesse* em colaborar é um evento externo que pode ser alavancado por incentivos da livraria ou pela cultura e conduta de colaboração em si.

Com o aumento da **Colaboração**, também tende a aumentar o volume de **Avaliações Humanas** na forma de revisões textuais, tanto **Revisões positivas**, quanto **Revisões negativas** (dada a heterogeneidade de opiniões de usuários). O peso de revisões positivas e negativas é ponderado por sua utilidade atribuída pelos usuários. **Revisões Negativas** com utilidade alta atribuída têm impacto negativo sobre as aquisições. Por outro lado, **Revisões Positivas** com utilidade alta aumentam a predisposição às aquisições.

4 DISCUSSÕES E CONSIDERAÇÕES FINAIS

A abordagem adotada neste trabalho para modelagem do sistema de recomendação baseou-se na visão sistêmica e no modelo CISM de Mario Bunge, o que propiciou uma análise qualitativa focada no nível de corte estabelecido para os componentes, elementos ambientais, as ligações entre estes e os mecanismos conjecturados. Por meio de diagramas de Boudon-Coleman e de diagramas de causa-efeito foi possível descrever e analisar o comportamento que decorre das interações sistêmicas no sistema de recomendações modelado. Juntamente com o conhecimento do domínio de sistemas de recomendação, a visão integrada das interligações e interações entre os componentes do sistema em seu contexto facilitou o delineamento dos processos que necessitam correção e aprimoramento.

As características de sistema sociotecnológico de recomendação para livraria virtual são representadas notadamente pela filtragem híbrida e pela contribuição humana por meio de avaliações em geral. Para funcionamento do sistema modelado e analisado, agentes humanos e agentes de software agem cooperativamente em função do artefato essencial (livro) para produção as recomendações. A incorporação de agentes artificiais com capacidade semântica e capazes de ligar dados obtidos de fontes diversas (*web*, redes sociais, repositórios públicos, dentre outras) pode proporcionar recomendações mais adequadas ao usuário. A análise automatizada de sentimentos e opiniões nos comentários textuais, assim como aferição da veracidade nos ratings expostos ao usuário pode equilibrar recomendações com informações adicionais confiáveis na decisão de aquisição dos livros. A contribuição humana para atribuição de *ratings* e a conduta de colaboração poderiam ser estimuladas por meio de incentivos personalizados para o perfil de cada usuário. Estas ações envolvem, porém, em uma maior extração de informações para consolidação destes perfis.

As filtragens de informações devem possibilitar um balanço entre acurácia nos resultados, diversidade de opções e novidade nas recomendações apresentadas ao usuário. Dados de contexto temporal, de localização e de circunstâncias pessoais sobre o usuário podem gerar recomendações mais acuradas, mas implicam em maior coleta e integração de informações sobre este. Diante destes aspectos, nota-se a importância do investimento na implantação e gestão de recursos de privacidade e de segurança de dados coletados, armazenados e transacionados para cultivar a participação ativa do usuário no sistema de recomendações.

REFERÊNCIAS

- ADOMAVICIUS, G.; TUZHILIN, A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. **IEEE Transactions on Knowledge and Data Engineering**, v. 17, n. 6, p. 734-749, 2005.
- BELLOGÍN, A.; CANTADOR, I.; CASTELLS, P. A comparative study of heterogeneous item recommendations in social systems. **Information Sciences**, v. 221, p. 142-169, 2013.
- BOBADILLA, J.; ORTEGA, F.; HERNANDO, A.; GUTIÉRREZ, A. Recommender systems survey. **Knowledge Based Systems**, 46, 109–132, 2013.
- BUNGE, M. **Chasing reality**: strife over realism. Toronto: University of Toronto Press, 2006
- BUNGE, M. **Emergence and convergence**: Qualitative novelty and the unity of knowledge. Toronto: University of Toronto Press, 2003.
- BUNGE, M. How does it work? The search for explanatory mechanisms. **Philosophy of the Social Sciences**, v. 34, n. 2, p. 182-210, 2004.
- BURKE, R. Hybrid recommender systems: Survey and experiments. **User Modeling and User Adapted Interaction**, v.12, n. 6, p. 331–370, 2002.
- ECCLES, D. W.; GROTH, P. T. Agent coordination and communication in sociotechnological systems: Design and measurement issues. **Interacting with Computers**, v. 18, n. 6, p. 1170-1185, 2006.
- FUCHS, C. The internet as a self-organizing socio-technological system. **Cybernetics and Human Knowing**, v. 12, n. 3, p. 57-81, 2005.
- JENNINGS, N.; SYCARA, K.; WOOLDRIDGE, M. A roadmap of agent research and development. **Autonomous Agents and Multi-Agent Systems**, v. 1, n. 1, p. 7-38, 1998.
- KERN, V. M. O sistemismo de Bunge: fundamentos, abordagem metodológica e aplicação a sistemas de informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 12., 2011, Brasília. **Anais...** Brasília: Thesaurus, 2011. p. 2693-2709.
- LI, Q.; WANG, J.; CHEN, Y. P.; LIN, Z. User comments for news recommendation in forum-based social media. **Information Sciences**, v. 180, n. 24, p. 4929-4939, 2010.
- SARWAR, B. et al. Analysis of recommendation algorithms for e-commerce. In: **Proceedings of the 2nd ACM Conference on Electronic Commerce**. Minneapolis, MN: ACM, 2000. p. 158-167.
- STERMAN, J. **Business dynamics**: Systems thinking and modeling for a complex world. Irwin McGraw-Hill, 2000.