

# **Gestão da informação em biblioteca universitária: uma proposta utilizando regras de associação na disseminação das informações de novas aquisições bibliográficas**

**Maciel F. da Silva<sup>1</sup>, Cláudio Ratke<sup>2</sup>**

<sup>1</sup>Bacharel em Sistemas de Informação  
Universidade Regional de Blumenau (FURB) – Blumenau, SC – Brasil

<sup>2</sup>Departamento de Sistemas e Computação  
Universidade Regional de Blumenau (FURB) – Blumenau, SC – Brasil  
maciel.f@gmail.com, ratke@inf.ufsc.br

***Resumo.** Este trabalho foi desenvolvido utilizando mineração de dados na disseminação de informação sobre novas aquisições do acervo bibliográfico de uma biblioteca universitária. O sistema proposto permite enviar, por e-mail, informações sobre as novas aquisições de acordo com as áreas selecionadas pelo usuário em seu perfil. A técnica de regras de associação foi empregada como um diferencial, objetivando enviar também informações de novas aquisições de outras áreas que podem interessar ao usuário, a partir das correlações detectadas pelo emprego desta. O sistema pode auxiliar nos serviços oferecidos pelas bibliotecas universitárias principalmente em relação às necessidades informacionais de seus usuários.*

## **1. Introdução**

Na sociedade contemporânea a informação assume um valor estratégico e as organizações focam na gestão da informação, utilizando as tecnologias, ferramentas e técnicas mais avançadas para obter, selecionar, mapear, organizar e disseminar a informação. Por isso, os sistemas de informação e de banco de dados atrelados aos rápidos avanços tecnológicos e as exigências competitivas do mercado tornam-se indispensáveis nas atividades das organizações.

As bibliotecas são organizações que tem a informação como foco de suas atividades e possuem um importante papel dentro das universidades. Na biblioteca universitária a disseminação da informação contribui para a produção do conhecimento no meio acadêmico, incentivando o uso adequado das informações e dos conhecimentos e minimizando os esforços dos usuários (DIAS, 2005).

A disseminação da informação é um importante veículo de comunicação entre a biblioteca e seus usuários, que têm necessidades informacionais distintas de acordo com o perfil e área de atuação. É papel da biblioteca fornecer ao seu usuário o produto informacional que supra a suas necessidades específicas, quer seja para estudo, para pesquisa ou para a tomada de decisão (TARAPANOFF; ARAÚJO JÚNIOR; CORMIER, 2000).

As tecnologias de informação são aliadas na disponibilização dos serviços de informação. Os usuários necessitam obter informações mais selecionadas em menor

tempo e para atendê-los com maior agilidade e eficiência as bibliotecas precisam utilizar as tecnologias disponíveis, tanto para a divulgação dos serviços oferecidos como para disponibilização de informações de acordo com o interesse dos seus usuários.

Ainda com o intuito de agregar valor na disseminação das informações sobre as novas aquisições de uma biblioteca universitária foram utilizadas tecnologias como a mineração de dados, em especial a técnica de regras de associação. Tal técnica é empregada para detectar informações de correlações ou tendências entre dados de transações em bases de dados, de modo que possam ser úteis na disponibilização de serviços de informação.

O principal objetivo neste trabalho foi promover a disseminação de informações referentes às novas aquisições de livros em uma biblioteca universitária. Em geral as bibliotecas adotam padrões internacionais de catalogação e classificação para os materiais bibliográficos, como por exemplo, a classificação Decimal de Dewey (CDD) que divide o conhecimento em áreas onde cada assunto é representado por um número que é atribuído as obras de acordo com o assunto destas.

O sistema proposto permite enviar periodicamente informações de livros para os usuários a partir de um perfil previamente definido com áreas de interesse, incluindo sugestões de livros de outras áreas, detectadas pelo uso das regras de associação, que também possam interessar.

## **2. Gestão e Disseminação da Informação Associada à Mineração de Dados**

A informação está inserida em todos os ambientes e em todas as atividades humanas, sociais, científicas, tecnológicas, culturais, políticas e econômicas, assumindo um novo status e importância na sociedade atual. Tais informações são variadas e produzidas de forma contínua, necessitando ser recuperadas, classificadas, organizadas, processadas, analisadas e difundidas pelas organizações no menor tempo possível (STAREC, 2005).

Por isso gerenciar a informação é primordial, e pode se tornar uma vantagem competitiva. Silva e Tomaél (2007, p. 1) afirmam que “para ser utilizada estrategicamente, é fundamental que a informação seja gerida em favor da sobrevivência e competitividade organizacional.”

A gestão da informação engloba a prospecção, seleção e obtenção da informação; o mapeamento e reconhecimento dos fluxos formais de informação; o tratamento, análise e armazenamento da informação utilizando tecnologias de informação; a disseminação e mediação da informação ao público interessado, e a criação e disponibilização de produtos e serviços de informação (VALENTIM, 2002).

No contexto acadêmico as bibliotecas são as principais responsáveis pela gestão da informação e possuem um papel importante dentro da universidade ao disponibilizar diversos tipos de materiais bibliográficos que proporcionam informações aos seus usuários. As bibliotecas universitárias “têm como missão oferecer aos seus usuários informações relevantes para a realização de pesquisas e para o ensino, procurando tornar o acesso, a recuperação e a localização das informações compatíveis com as suas necessidades.” (LUCAS; SOUZA, 2007, p. 2).

A disseminação da informação é uma importante etapa da gestão da informação, completando o processo de seleção, tratamento e análise da informação. Disseminar significa difundir, propagar, semear ou espalhar por muitas partes (MICHAELIS, 1998).

As bibliotecas universitárias devem oferecer serviços de informação que satisfaçam as expectativas e as necessidades dos usuários (CORTES; LOPES, 2008). Com um volume cada vez maior de informações disponíveis é fundamental que a informação enviada a um usuário seja o mais próximo daquilo que ele necessita.

Por isso, oferecer um serviço de Disseminação Seletiva da Informação (DSI), que fornece ao usuário uma relação periódica de fontes informacionais, relacionadas com sua área e detectadas a partir do preenchimento de um perfil de interesse (SOUTO; PORTELA, 2003/2004) torna-se um diferencial para as bibliotecas.

No caso específico das bibliotecas universitárias os usuários têm necessidades informacionais distintas de acordo com o perfil de cada um e de sua área de atuação. Esses usuários também necessitam estar sintonizados com as publicações de seu interesse e, principalmente, com as novas aquisições efetuadas pela biblioteca.

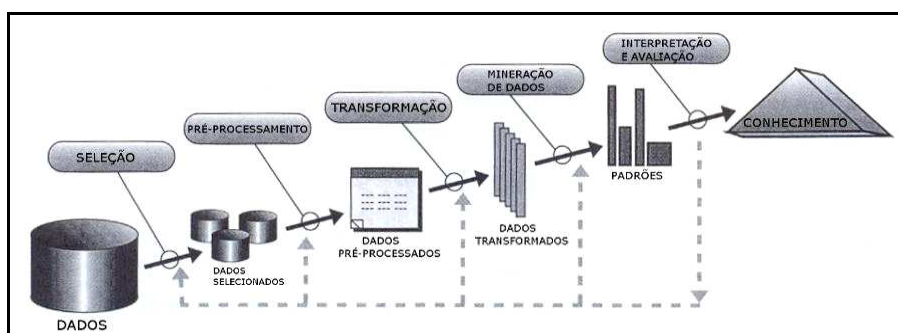
No cenário atual, em que as pessoas têm cada vez menos tempo, os serviços online são os aliados para atender essas demandas, pois com o uso da Internet não existem barreiras de tempo e espaço, e o serviço é amplamente acessível. Os usuários “para atender às exigências da sociedade atual, estão cada vez mais à procura de praticidade e rapidez, vantagens estas, que as inovações tecnológicas proporcionam.” (CORTES; LOPES, 2008, p.117). E a “[...] Internet trouxe um aumento considerável quanto à oferta de informações e às possibilidades de sua disseminação.” (DIAS, 2005, p.86).

Além disso, com a necessidade de obter informações cada vez mais rápidas surgem novas metodologias utilizadas para a análise qualitativa e quantitativa, como a mineração de dados, que é uma tecnologia muito útil no gerenciamento de informação. A mineração de dados ou a extração de dados específicos de enormes campos de informação é um conceito em moda na gestão da informação (FERAUD, 2004).

## **2.1. Descoberta de Conhecimentos em Banco de Dados**

Nos últimos anos, os avanços tecnológicos têm proporcionado a retenção de grandes volumes de dados para as empresas, atraindo a atenção para a mineração de dados, devido à necessidade de transformar esses dados em conhecimentos úteis (HAN; KAMBER, 2006). Da necessidade de técnicas e ferramentas específicas para análise e extração de conhecimento surge o *Knowledge Discovery in Databases* (KDD) ou descoberta de conhecimento em base de dados, visto como “um processo não trivial de identificação de padrões válidos, novos (desconhecidos até então), potencialmente úteis, e compreensíveis.” (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996, p. 6, tradução nossa).

Assim, o processo de descoberta de conhecimento envolve várias etapas não lineares, onde o processo envolve interação e a iteração significativa, que pode conter *loops* entre duas etapas ou mais, sendo a mineração de dados uma das etapas do KDD (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996). A Figura 1 apresenta as etapas do KDD.



Fonte: adaptado de Fayyad, Piatetsky-Shapiro e Smyth (1996, p.10).

**Figura 1. Etapas do Processo da Descoberta de Conhecimento em Banco de Dados.**

Resumidamente as etapas do KDD são definidas em:

- a) dados: compreender o domínio da aplicação e os objetivos do usuário final;
- b) seleção: selecionar o conjunto de dados para análise na qual a descoberta deve ser realizada;
- c) pré-processamento: tratar ruídos, como campos de dados faltantes, nulos ou repetidos, corrigir prováveis erros relevantes para a mineração dos dados;
- d) transformação: formatar os dados conforme o objetivo da tarefa, utilizando redução de dimensionalidade ou métodos de transformação para reduzir o número de variáveis em consideração;
- e) mineração de dados: definir o método e a técnica a ser utilizada para a descoberta de padrões, bem como o algoritmo que irá realizar a extração desses padrões de interesse, ajustando-o se necessário;
- f) interpretação e avaliação: analisar os resultados obtidos através da mineração, com possibilidade de retorno aos passos anteriores;
- g) conhecimento: incorporar o conhecimento obtido ao sistema, documentar e reportar os resultados às partes interessadas (FAYYAD; PIATETSKY-SHAPIRO; SMYTH, 1996).

## 2.2. Regras de Associação

Dentre as técnicas da mineração de dados, a regra de associação se destaca por ser uma técnica que permite encontrar padrões, associações ou correlações em conjuntos de itens frequentes de uma base de dados transacional, relacional ou de outros tipos de repositórios de informação (NEVES, 2002).

Um exemplo comum da utilização de regras de associação para mineração na descoberta de padrões através das compras realizadas, transações feitas, por clientes em supermercados ou lojas, na qual uma mercadoria está relacionada a outra, “[...] conhecida também como transações de cestas de compras.” (TAN; STEINBACH; KUMAR, 2009, p. 389). A partir desses padrões encontrados, o estabelecimento poderia realizar promoções de vendas em determinados produtos ou também conhecer o perfil de seus consumidores.

Regras de associação basicamente têm duas medidas utilizadas para análise, o suporte e a confiança. “O suporte determina a frequência na qual uma regra é aplicável a um determinado conjunto de dados, e a confiança determina a frequência na qual os itens em  $Y$  aparecem em transações que contenham  $X$ .” (TAN; STEINBACH; KUMAR, 2009, p. 392).

Formalizando o problema para mineração de regras de associação conforme Agrawal et al (1996), seja  $I = \{i_1, i_2, \dots, i_m\}$  um conjunto de itens e  $D$  um conjunto de transações, onde cada transação  $T$  é um conjunto de itens (*itemset*) tal que  $T \subset I$ . Associado a cada transação está um único identificador, chamado identificador único de transação (TID). Assim, uma regra de associação é uma implicação na forma expressa por  $X \Rightarrow Y$  ( $X$  então  $Y$ ), sendo  $X$  denominado de antecedente e  $Y$  conseqüente da regra, onde  $X$  e  $Y$  são *itemsets* distintos, representados na expressão  $X \subset I, Y \subset I$  e  $X \cap Y = \emptyset$ .

O suporte da regra  $X \Rightarrow Y$  é definido na expressão como  $sup(X \Rightarrow Y) = \sigma(X \cup Y) / D$ , ou seja, é a razão entre o contador de suporte contendo  $X$  e  $Y$  e o total de transações. E a confiança da regra  $X \Rightarrow Y$  é definida na expressão como  $conf(X \Rightarrow Y) = \sigma(X \cup Y) / \sigma(X)$ , é a razão entre o contador do suporte contendo  $X$  e  $Y$  e o contador do suporte contendo  $X$  (TAN; STEINBACH; KUMAR, 2009).

Para que uma regra seja interessante, ou dita forte, deve atender os critérios, tenha suporte  $\geq$  minsup e a confiança  $\geq$  minconf, sendo minsup e minconf os limites de suporte e confiança respectivamente definidos (TAN; STEINBACH; KUMAR, 2009).

Então, a estratégia utilizada é decompor o problema em duas etapas, encontrar todos os conjuntos de itens que tenham o seu suporte maior que o suporte mínimo estabelecido, estes conjuntos são chamados de conjuntos de itens frequentes. Após encontrar os conjuntos de itens frequentes, gerar as regras que atendam a um mínimo de confiança (AGRAWAL et al., 1996).

O algoritmo Apriori, muito conhecido por encontrar regras de associação, foi um dos primeiros a adotar a abordagem de suporte mínimo estabelecido para gerar conjuntos de itens candidatos (TAN; STEINBACH; KUMAR, 2009). Basicamente é dividido em duas fases e funciona iterativamente de modo a encontrar conjuntos de itens frequentes em uma determinada base de dados para que posteriormente sejam extraídas as regras de associação.

Na primeira fase, inicia com todos os itens da base como sendo um conjunto de um item, é feita a contagem do suporte para cada conjunto, são eliminados os conjuntos que não atendem ao suporte mínimo definido, restando somente os conjuntos de um item ditos frequentes. A partir desse conjunto, são gerados novos conjuntos candidatos, contendo agora dois itens cada conjunto, verifica-se novamente o suporte e em seguida são eliminados novamente aqueles, cujo suporte é inferior ao mínimo definido. Assim, o processo é repetido  $n$  vezes até não haver mais conjuntos de itens candidatos, resultando nos conjuntos contendo os itens mais frequentes da base de dados (TAN; STEINBACH; KUMAR, 2009).

Encontrados então os conjuntos de itens frequentes, na segunda fase são geradas as regras de associação candidatas e extraídas as regras fortes que atendam a um suporte mínimo de confiança utilizado como parâmetro.

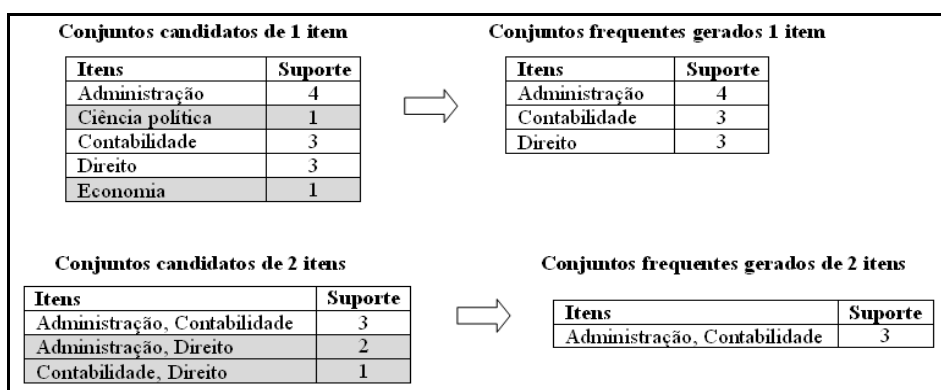
Sendo assim, seja uma base de dados de transações qualquer, onde os itens são algumas áreas (Quadro 1), deseja-se encontrar itens frequentes para extrair regras, cujo valor do suporte mínimo deve ser de 50% ou mais, (contador do suporte maior ou igual a 3), e a confiança mínima da regra 70% ou superior.

**Quadro 1. Exemplo de uma base de dados com transações**

TID	Itens
1	Administração, Contabilidade
2	Direito, Administração, Contabilidade
3	Administração, Economia, Direito
4	Administração, Contabilidade
5	Direito, Ciência política

São gerados os conjuntos candidatos e encontrados os conjuntos de itens frequentes dessa base, conforme Quadro 2.

**Quadro 2. Geração dos conjuntos de itens frequentes**



Com base nos conjuntos de itens frequentes gerados, são extraídas então as regras de associação que possuem relação com outros itens e que satisfaçam a confiança mínima definida para o problema em questão, de acordo com o Quadro 3.

**Quadro 3. Exemplo da extração de regras de associação com base no algoritmo Apriori.**

Conjuntos de itens frequentes gerados	
Conjuntos	Suporte
Administração	4 (80%)
Contabilidade	3 (60%)
Direito	3 (60%)
Administração, Contabilidade	3 (60%)

Regras de associação geradas	
Regras geradas	Confiança
Administração => Contabilidade	$3 / 4 = 0,75$ (75%)
Contabilidade => Administração	$3 / 3 = 1$ (100%)

Das duas regras de associação geradas, podemos verificar que com relação à base de dados mencionada, a primeira regra mostra que 75% das transações onde há Administração, encontramos Contabilidade. Já na outra regra, o nível da confiança é ainda maior, pois 100% das transações contendo Contabilidade também há Administração.

Assim podemos perceber a correlação entre os itens de uma determinada base de dados contendo dados de transações, e a incidência em que um item pode apresentar em relação a outro, sugerindo então informações de tendências que podem ser aproveitadas nas organizações de diversas formas.

### 3. Desenvolvimento do Sistema

O sistema proposto compreende em um portal disponibilizado na internet, desenvolvido em linguagem de programação *PHP* com *HTML* e *CSS*, hospedado em um servidor *Apache*. E ainda um módulo gerencial para promover a disseminação das informações aos usuários, desenvolvido em ambiente de programação *Delphi*, e alocado em um servidor com plataforma *Windows Server*. O gerenciador de banco de dados utilizado foi o *Oracle 10g* e *PL/SQL* para criação de funções auxiliares do banco.

### 4. Implementações e Resultados

Para o usuário acessar ou configurar seu perfil no portal, deve-se efetuar o *login* com seu nome de usuário e sua senha. Ao se autenticar é apresentado ao usuário o seu perfil, no qual ele irá selecionar as suas preferências. No topo da página são mostradas algumas opções, como por exemplo, alteração do *e-mail*, menu de ajuda ou sair da página. No centro da página encontram-se as áreas ou assuntos de interesse, disponibilizadas em um total de cem áreas, na qual o usuário pode optar selecionando cada uma delas. A Figura 2 apresenta parte da página principal do perfil do usuário.

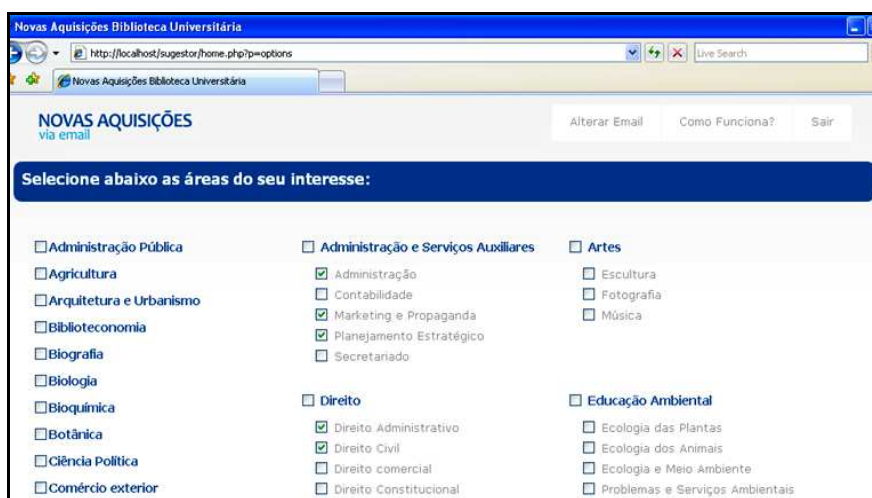


Figura 2. Parte da página principal de preferências do usuário

Após o usuário selecionar as áreas de sua preferência, mais ao final da página o usuário decide por autorizar ou não o recebimento de informações em seu *e-mail*. Por fim, irá clicar no botão “Salvar” para gravar as suas informações, conforme Figura 3.

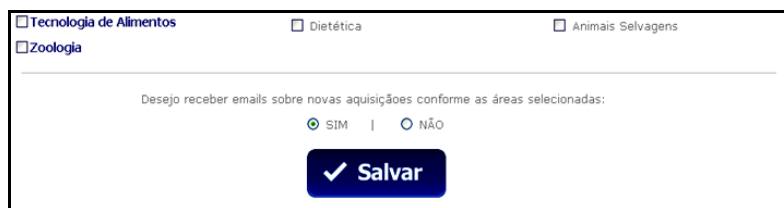


Figura 3. Parte da página principal para autorização e salvar preferências

A disseminação das informações é feita por um aplicativo que é acessado somente pelo administrador, e é executada periodicamente de forma automatizada, permitindo a geração e o envio dos relatórios personalizados para cada usuário. O administrador ainda poderá efetuar a execução manual do aplicativo caso seja necessário, e também efetuar alterações de parâmetros, consultas e relatórios estatísticos de utilização do sistema. Entretanto, a maior importância desse módulo encontra-se nas sub-rotinas de processamento das informações que serão encaminhadas aos usuários e juntamente com a mineração das regras de associação.

Basicamente a ferramenta efetua a busca por novas aquisições, correspondente a cada área que o usuário selecionou em seu perfil. Em seguida, são extraídas as áreas correlatas do perfil do usuário, geradas a partir das regras de associação, e então é feita uma nova busca por novas aquisições conforme cada área correlata gerando as sugestões que possam ser de interesse ao usuário.

É importante ressaltar que as áreas correlatas que geram as sugestões podem diferir das áreas compreendidas no site, visto que estariam disponíveis apenas as principais áreas, já previamente estipuladas pela instituição, entretanto o usuário não ficaria privado de estar recebendo informações de outras áreas e que poderiam ser de relevância.

Por fim, se houver alguma informação de nova aquisição no relatório gerado, tanto informações conforme as áreas selecionadas no perfil, como sugestões detectadas a partir das áreas correlatas, são enviadas para o *e-mail* do respectivo usuário, a exemplo da Figura 4.

**NOVAS AQUISIÇÕES CONFORME PERFIL SELECIONADO:**

**Administração:**  
[Gestão de idéias para inovação contínua. José Carlos Barbieri, Antonio Carlos Teixeira Álvares, Jorge Emanuel Reis Cajazeira.](#)  
[Uma aula com Drucker: as lições do maior mestre de administração. William Cohen ; tradução Afonso Celso da Cunha Serra.](#)

**Contabilidade:**  
[Contabilidade geral: introdução à contabilidade societária. Natan Szuster ... \[et al.\].](#)

**Marketing e Propaganda:**  
[Comunicação integrada de marketing: uma visão global. Roberto Correa.](#)  
[Fundamentos estratégicos de marketing. Edson Roberto Scharf, Vanessa Edy Dagnoni Mondini.](#)  
[Macro marketing: uma perspectiva social. Reed Moyer; tradução de Ebreia de Castro Alves.](#)

---

**LIVROS SUGERIDOS PELO SISTEMA:**

[Uma aula com Drucker: as lições do maior mestre de administração. William Cohen ; tradução Afonso Celso da Cunha Serra.](#)  
[Metodologia científica e educação. Agripa Faria Alexandre.](#)  
[Metodologia de pesquisa em ciência da computação. Raul Sidnei Wazlawick.](#)  
[Pesquisa e monografia jurídica na era da informática. Liliane dos Santos Vieira.](#)

Esta mensagem foi gerada automaticamente.  
Para não receber mais estas informações, desabilite a autorização em seu perfil.

**Figura 4. Parte do e-mail com relatório personalizado conforme perfil**

Para gerar as sugestões foi feita uma adaptação de Apriori no caso específico da biblioteca. Enquanto Apriori encontra primeiramente os conjuntos de itens frequentes na base de transações e posteriormente gera as regras a partir dos itens frequentes, neste trabalho, foi necessário encontrar as áreas correlatas a partir de cada área disponibilizada no perfil do usuário. Assim, as áreas do perfil foram tratadas como itens frequentes e consideradas antecedentes da regra.

Depois de calculado o suporte da área que antecede a regra, são encontradas as áreas correlatas, os consequentes da regra, que satisfazem o suporte de confiança mínimo definido previamente pelo administrador. Devido à grande quantidade de áreas



na qual os livros estão classificados, percebeu-se que os percentuais de confiança na grande maioria das regras geradas não ultrapassaram os 50%. Já para um percentual abaixo de 20% muitas regras são encontradas nas quais muitas informações de livros são geradas, podendo conter informações em excesso ao usuário. Assim, a sugestão de um percentual mínimo ideal no contexto verificado estaria compreendida entre 20% e 45% para a geração de regras, definido a critério do administrador.

O modelo de regra de associação adotada para este trabalho diz-se do primeiro nível, sendo apenas uma área para o antecedente e uma área para o consequente da regra gerada, por exemplo, a área *X* do perfil possui a área *Y* como consequente.

O identificador único de transação (TID) utilizado para a extração das regras, a partir das transações de livros emprestados, é o código do usuário. É importante ressaltar que o foco deste trabalho está na área que o livro pertence, conforme o número de classificação da CDD. Assim, independente da quantidade de livros ou títulos emprestados pelo usuário, cada área é contabilizada uma única vez por usuário. Um exemplo dessa abordagem está representado no Quadro 4.

**Quadro 4. Exemplo da base de dados de transações da abordagem adotada**

TID (cód. usuário)	Áreas (cód. área da CDD)					
	005.1	342.1	630	658	370	813
001	1	0	0	1	0	1
002	0	1	0	0	1	0
003	0	1	1	0	0	1
004	1	1	0	1	0	1
005	0	1	1	1	0	1
<b>Contador</b>	<b>2</b>	<b>4</b>	<b>2</b>	<b>3</b>	<b>1</b>	<b>4</b>

Para encontrar as áreas correlatas e gerar as regras de associação, utilizou-se dados reais de uma biblioteca universitária, baseando-se nos registros das transações do histórico de empréstimos realizados em um período de aproximadamente um ano. Foi necessário ainda efetuar um pré-processamento, removendo ruídos e reunindo somente os dados de interesse, e posteriormente alocando esses dados em uma nova tabela, do tipo *view* materializada, para então aplicar a mineração, pois havia mais de um milhão e meio de registros.

A tabela visão foi criada no intuito de agilizar e facilitar a extração das regras de associação, uma vez que os dados reunidos eram de várias tabelas incluindo a tabela de histórico de empréstimos, sendo assim a *view* deve ser atualizada periodicamente para que as regras estejam sempre em conformidade com a demanda mais atual dos empréstimos realizados. O Quadro 5 apresenta um fragmento da instrução que gera a *view*.

**Quadro 5. Fragmento da instrução que gera a view**

```
SELECT DISTINCT C.CD_PESSOA, COUNT(NR_REGISTRO)AS QTD, C.NR_REGISTRO, A.CD_MFN,
T.AREA AS CD_AREA, INFO_AREA(T.AREA) AS DS_AREA
FROM CIRCULACAO_HISTORICO_ANO_BASE C, ACERVO_DICIONARIO_MFN A,
(SELECT OBRA(CD_MFN,'082') AS AREA, CD_MFN
FROM ACERVO_DADOS_MFN
WHERE CD_MFN IN (SELECT CD_MFN FROM ACERVO_DADOS_MFN)) T
WHERE C.NR_REGISTRO = A.DS_TERM0
AND C.TP_OPERACAO = 'E'
AND C.NR_REGISTRO LIKE 'CG%'
AND A.CD_MFN = T.CD_MFN
GROUP BY C.CD_PESSOA, C.NR_REGISTRO, A.CD_MFN, T.AREA
```

Assim, a partir de uma área selecionada no perfil do usuário, busca-se na *view* quais são os usuários que fizeram empréstimos de livros daquela mesma área, e depois verifica-se quais são as outras áreas relacionadas por esses mesmos usuários, por exemplo, usuários que emprestam livros da área X, também emprestam livros da área Y, digamos que esta seja uma regra gerada  $X \Rightarrow Y$ , dessa maneira, um usuário que selecionou a área X em seu perfil, estaria recebendo informações da área X e também informações livros da área Y sob a forma de sugestões.

No Quadro 6, é apresentado um fragmento de código que gera as regras de associação ou áreas correlatas em relação às áreas selecionadas pelo usuário em seu perfil.

**Quadro 6. Fragmento de código para gerar as regras de associação**

```

qry_area_regra_assoc.Close;
qry_area_regra_assoc.SQL.Clear;
qry_area_regra_assoc.SQL.Add('SELECT RS.CD_AREA2');
qry_area_regra_assoc.SQL.Add('FROM (SELECT A.CD_AREA AS CD_AREA1, A.DS_AREA AS DS_AREA1, B.CD_AREA AS CD_AREA2,
B.DS_AREA AS DS_AREA2, COUNT(*) AS OTD_AREA2);
qry_area_regra_assoc.SQL.Add(' FROM (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA FROM TAB_PADRAO) B.);
qry_area_regra_assoc.SQL.Add(' (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA FROM TAB_PADRAO WHERE CD_AREA = ' + cd_area + ') A);
qry_area_regra_assoc.SQL.Add(' WHERE A.CD_PESSOA = B.CD_PESSOA);
qry_area_regra_assoc.SQL.Add(' AND A.CD_AREA <> B.CD_AREA);
qry_area_regra_assoc.SQL.Add(' GROUP BY A.CD_AREA, A.DS_AREA, B.CD_AREA, B.DS_AREA);
qry_area_regra_assoc.SQL.Add(' ORDER BY OTD_AREA2 DESC, B.DS_AREA) RS.);
qry_area_regra_assoc.SQL.Add(' (SELECT COUNT(*) AS TOTAL_AREA1);
qry_area_regra_assoc.SQL.Add(' FROM (SELECT DISTINCT CD_AREA, DS_AREA, CD_PESSOA);
qry_area_regra_assoc.SQL.Add(' FROM TAB_PADRAO WHERE CD_AREA = ' + cd_area + ') A) OTD, PARAMETRO_REGRA_ASSOC PR);
qry_area_regra_assoc.SQL.Add('WHERE (RS.OTD_AREA2 / OTD.TOTAL_AREA1) * 100 >= PR.PERCENT_SUPOORTE_CONF);
qry_area_regra_assoc.Open;

if not qry_area_regra_assoc.IsEmpty then
begin
  qry_area_regra_assoc.First;
  while not qry_area_regra_assoc.Eof do
  begin
    if cd_area_regra_assoc = "" then
    begin
      cd_area_regra_assoc := (" + qry_area_regra_assoc.FieldByName('CD_AREA2').AsString + "");
    end
    else
    begin
      cd_area_regra_assoc := (cd_area_regra_assoc + ' + ' + "" +
      qry_area_regra_assoc.FieldByName('CD_AREA2').AsString + "");
    end;
  end if
  qry_area_regra_assoc.Next;
end;
end while
end;
end if
qry_area_regra_assoc.Close;

```

A utilização da técnica para minerar as áreas correlatas a partir do perfil do usuário gerou resultados interessantes, revelando relações que até então eram desconhecidas entre algumas áreas. Dentre os casos de regras geradas para um percentual mínimo de confiança de 30%, podemos mencionar as áreas “Artes” e “Literatura Inglesa” como sendo os antecedentes das regras, por exemplo, teríamos as seguintes regras geradas pelo sistema, conforme o Quadro 7.

**Quadro 7. Exemplo de regras geradas pelo sistema**

Regras geradas	Confiança
Artes (700) => História da Moda (391)	42%
Artes (700) => Arte e sociedade (701)	37%
Artes (700) => Desenho industrial (745.2)	37%
Literatura Inglesa (823) => Literatura Americana (813)	58%
Literatura Inglesa (823) => Literatura Brasileira (869.93)	32%

Assim, o usuário que tivesse selecionado “Artes” e “Literatura Inglesa” em seu perfil, estaria recebendo além de livros dessas áreas, também receberia como sugestões livros classificados nas áreas, conforme a CDD, “História da Moda”, “Arte e Sociedade”, “Desenho industrial”, “Literatura Americana” e “Literatura Brasileira”.

## 5. Conclusão

Gerenciar a informação em uma biblioteca universitária, especificamente a disseminação da informação sobre novas aquisições do acervo bibliográfico, mostrou-se importante para a criação de um canal de comunicação eficaz com seus usuários, possibilitando a disseminação de informações sobre o acervo diretamente para o *e-mail* deles, minimizando tempo gasto com buscas e maximizando o uso dos recursos tecnológicos.

A utilização da tecnologia de mineração de dados para extração de informações em base de dados, especificamente da técnica de regra de associação, foi útil também como ferramenta gerencial, pois permitiu encontrar correlações entre algumas áreas que ainda eram desconhecidas e em outros casos possibilitou constatar associações que se acreditava existir.

O emprego das regras de associação para o desenvolvimento do sistema possibilitou prever possíveis necessidades de informação ao verificar que quem empresta obras de uma área também empresta de outra determinada área. Isto permitiu criar um sistema mais elaborado que além de enviar por *e-mail* as informações de interesse do usuário, conforme a seleção feita previamente por este, também envia informações de obras que podem interessar a estes usuários, detectadas pelo uso das regras de associação.

Entretanto, foram encontradas algumas limitações durante o desenvolvimento do sistema, como por exemplo, a quantidade restrita de áreas disponível na interface *web* para o usuário, áreas estas que já estavam predefinidas pela biblioteca como sendo as áreas mais gerais da CDD relacionadas aos cursos da universidade, e desconsiderando as mais específicas.

Por fim, o sistema proposto irá proporcionar que a biblioteca disponibilize um serviço de disseminação da informação mais condizente com as necessidades dos usuários. E o usuário se beneficia com a disponibilização do serviço, pois terá um instrumento de atualização na sua área de atuação.

## Referências

- AGRAWAL, Rakesh et al. Fast discovery of association rules. In: FAYYAD, Usama M. et al. **Advances in knowledge discovery and data mining**. Menlo Park: AAAI: MIT, 1996. p. 307-328.
- CORTES, Márcia Della Flora; LOPES, Marilisa Leite. As bibliotecas universitárias federais brasileiras e a acessibilidade das informações em seus websites. **Revista ACB: Biblioteconomia em Santa Catarina, Florianópolis**, v. 13, n.1, p.117-129, 2008. Disponível em: <<http://revista.acbsc.org.br/index.php/racb/article/viewArticle/552>>. Acesso em: 01 jun. 2010.
- DIAS, Simone Lopes. **A disseminação da informação mediada por novas tecnologias e a educação do usuário na Biblioteca Universitária**. 2005. 138 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005. Disponível em:

- <[http://polo1.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/dias\\_sl\\_me\\_mar.pdf](http://polo1.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/dias_sl_me_mar.pdf)>. Acesso em 01 jun. 2010.
- FAYYAD, Usama M.; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From data mining to knowledge discovery: an overview. In: FAYYAD, Usama M. et al. **Advances in knowledge discovery and data mining**. Menlo Park: AAAI: MIT, 1996. p. 1-34.
- FERAUD, Genevière. Um século de gestão da informação. In: DAVENPORT, Thomas H.; MARCHAND, Donald A.; DICKSON, TIM. **Dominando a gestão da informação**. Porto Alegre: Bookman, 2004. p. 39-44.
- HAN, Jiawei; KAMBER, Micheline. **Data Mining: concepts and techniques**. 2nd ed. Amsterdam: Elsevier, 2006.
- LUCAS, Elaine R. de Oliveira; SOUZA, Nicole Amboni de. Disseminação seletiva da informação em Bibliotecas Universitárias sob o prisma do Customer Relationship Managment. **Informação & Informação**, Londrina, v. 12, n. 1, p. 1- 17, 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewArticle/1745>>. Acesso em: 01 jun. 2010.
- MICHAELIS: moderno dicionário da língua portuguesa. 4. ed. São Paulo : Melhoramentos, 1998.
- NEVES, João Manuel Poças Marques das. **Ambiente de pós-processamento para regras de associação**. 2002. 118 f. Dissertação (Mestrado em Análise de Dados e Sistemas de Apoio à Decisão) - Faculdade de Economia, Universidade do Porto, Porto, 2002. Disponível em: <[http://www.fep.up.pt/cursos/mestrados/madsad/teses/tese\\_mestrado\\_joao\\_pocas\\_pos\\_processamento\\_ra.pdf](http://www.fep.up.pt/cursos/mestrados/madsad/teses/tese_mestrado_joao_pocas_pos_processamento_ra.pdf)>. Acesso em: 30 maio 2010.
- SILVA, Terezinha Elisabeth da; TOMAÉL, Maria Inês. A gestão da informação nas organizações. **Informação e Informação**, Londrina, v. 12, n. 2, p. 1-2, jul./dez. 2007. Disponível em: <<http://www.uel.br/revistas/uel/index.php/informacao/article/viewFile/1806/1540>>. Acesso em: 01 jun. 2010.
- SOUTO, Leonardo Fernandes; PORTELA, Patrícia de Oliveira. O SDI como instrumento de educação continuada: a responsabilidade das universidades no treinamento dos usuários. **Revista da ACB: Biblioteconomia em Santa Catarina**, Florianópolis, v. 8/9, p. 123-133, 2003/2004. Disponível em: <<http://revista.acbsc.org.br/index.php/racb/article/view/411>>. Acesso em 05 mar. 2010.
- STAREC, Claudio. A dinâmica da informação: a gestão estratégica da informação para a tomada de decisão nas organizações. In: STAREC, Claudio; GOMES, Elisabeth; BEZERRA, Jorge. **Gestão estratégica da informação e inteligência competitiva**. São Paulo: Saraíva, 2005. p. 47-64.
- TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao data mining: mineração de dados**. Rio de Janeiro: Ciência Moderna, 2009.

TARAPANOFF, Kira; ARAÚJO JÚNIOR, Rogério Henrique de; CORMIER, Patricia Marie Jeanne. Sociedade da informação e inteligência em unidades de informação. **Ciência da Informação**, Brasília, v. 29, n.3, p. 91-100, set./dez. 2000. Disponível em: <[http://www.scielo.br/scielo.php?pid=S0100-19652000000300009&script=sci\\_arttext&tlng=es](http://www.scielo.br/scielo.php?pid=S0100-19652000000300009&script=sci_arttext&tlng=es)>. Acesso em: 30 maio 2010.

VALENTIM, Marta Ligia Pomim. Inteligência Competitiva em Organizações: dado, informação e conhecimento. **DataGramZero**: Revista de Ciência da Informação, Rio de Janeiro,- v.3, n.4, p.1-13, ago. 2002. Disponível em: <[http://www.dgz.org.br/ago02/Art\\_02.htm](http://www.dgz.org.br/ago02/Art_02.htm)>. Acesso em: 16 maio 2010.