

Descoberta de padrões em bases de dados utilizando Técnicas Adaptativas

A. H. Tchemra, R. Camargo

Resumo – Este artigo tem como objetivo apresentar como os métodos da Tecnologia Adaptativa podem ser aplicados, como formas alternativas de solução, em ferramentas de apoio a decisão. O método proposto trata a Mineração de Dados Adaptativa, que consiste de um processo de descoberta de padrões e técnicas de mineração de dados agregadas ao conceito de adaptatividade. O estudo da aplicabilidade dessa proposta é desenvolvido a partir de uma base de dados organizada de forma multidimensional. Nesta aplicação, a tomada de decisão utiliza os padrões descobertos, que representam os critérios de um problema de decisão, como entrada numa tabela de decisão adaptativa baseada em métodos multicritério e técnicas adaptativas.

Palavras-chave: tecnologia adaptativa, mineração de dados adaptativa, tabela de decisão adaptativa, tomada de decisão, múltiplos critérios, mineração de dados.

I. INTRODUÇÃO

A área de Tecnologia Adaptativa pesquisa métodos adaptativos que permitem aos dispositivos adaptativos alterar suas estruturas internas e seus comportamentos de forma autônoma, de acordo com as necessidades requeridas durante o processo para obtenção da solução de determinados problemas [1].

O objetivo deste artigo é apresentar um estudo de um modelo de Mineração de Dados Adaptativa, como ferramenta para descoberta de padrões para auxiliar no processo de tomada de decisão.

Esse estudo trata a Mineração de Dados Adaptativa como um processo de identificação de tendências ou padrões em um conjunto de dados para orientar decisões a serem tomadas, por exemplo, dentro de atividades de planejamento estratégico.

Conforme [2], quando os padrões são definidos através de ferramentas de Mineração de Dados, é possível que as informações existentes e as inesperadas possam ser melhor investigadas.

De acordo com [2], Mineração de Dados investiga padrões ocultos existentes em grandes bases de dados. Essa ferramenta combina a identificação de padrões no conjunto de dados com tecnologias avançadas de aprendizagem computacional, envolvendo técnicas como inteligência artificial e estatística, descobrindo relações não explícitas. Os dados são, em geral, obtidos a partir de sistemas transacionais existentes, filtrados e armazenados

em um depósito único de dados. Atualmente, observa-se que o avanço tecnológico e a redução de custo de armazenamento tornaram a investigação de dados corporativos uma realidade.

Conforme [1], as características da Tecnologia Adaptativa conferem aos seus métodos a classificação de sistemas inteligentes, permitindo desta forma um tratamento alternativo para a mineração de dados com a implementação de técnicas adaptativas, objetivando um estudo de suas características em relação aos métodos existentes.

Este artigo apresenta na seção II os conceitos principais sobre Mineração de Dados, seguida pela seção III que descreve o modelo de Mineração de Dados Adaptativa proposto. A seção IV mostra a forma como os resultados obtidos na Mineração de Dados Adaptativa podem ser utilizados em um processo de tomada de decisão, com o auxílio de uma tabela de decisão adaptativa. Finalmente, na seção V são apresentadas as conclusões e considerações sobre o modelo adaptativo proposto.

II. CONCEITO DE MINERAÇÃO DE DADOS

Os estudos sobre Mineração de Dados ou Data Mining tiveram origem em análise estatística na década de 60, que evoluíram, posteriormente nos anos 80, para novas técnicas de inteligência artificial, tais como lógica fuzzy, redes neurais, árvores de decisão [4].

Mineração de dados, segundo [5], consiste em um conjunto de técnicas utilizadas na exploração de conjuntos de dados, normalmente mantidos em tabelas, formando um banco de dados. A mineração de dados tem como objetivo o descobrimento de relacionamentos complexos envolvendo conceitos, tais como: padrões, regras, fatos em dados armazenados.

Conforme [2], o processo de descoberta de conhecimento e mineração de dados (KDD, *Knowledge Discovery and Data Mining*), pode ser tratado em quatro etapas:

1. Seleção de dados: etapa para determinar o agrupamento de dados e atributos de interesse;
2. Limpeza dos dados: consiste na remoção de ruídos, na transformação de alguns campos e na criação de campos combinados;
3. Mineração dos dados: aplicação de algoritmos específicos para extrair padrões de interesse;
4. Avaliação: etapa em que os padrões descobertos são disponibilizados para os usuários em forma inteligível, facilitando a visualização.

Deve-se ressaltar que o conceito de padrão, segundo

[2], é uma unidade de informação ou atributo de um registro que se repete, ou então é uma seqüência de informações/atributos presentes em uma estrutura que se repete.

Segundo [3], existem cinco tipos de técnicas usadas para a Mineração de Dados:

1. **Associações:** que identificam afinidades entre um conjunto de dados em um grupo de registros; por exemplo: 72% de todos os registros que contêm itens A, B e C, também contêm itens D e E; dessa maneira, regras associativas procuram estabelecer ligações entre um elemento e outro;
2. **Padrões Seqüenciais:** que identificam seqüências de registros que ocorrem em decorrência de outros. Por exemplo: na ocorrência de um evento A, 32% dos clientes com determinadas características realizarão o evento B, em um determinado espaço de tempo;
3. **Classificação:** que divide as classes predefinidas, permitindo que registros de uma classe permaneçam próximos. Exemplificando: poderia haver classes de registros quanto à freqüência do comparecimento de clientes em uma agência bancária: infreqüentes (nunca freqüentam a agência), freqüentes (comparecem de modo freqüente) e ocasionais (ocasionalmente freqüentam a agência);
4. **Agrupamento:** a partir da base de dados, descobre classes ocultas, enquanto que a classificação já inicia com classes predefinidas;
5. **Previsão:** que tem como objetivo o cálculo de previsão do valor futuro de uma variável, como por exemplo, prever uma determinada projeção de vendas, considerando registros devidamente classificados.

III. PROPOSTA DE UM MODELO DE MINERAÇÃO DE DADOS ADAPTATIVA

As pesquisas sobre técnicas adaptativas e mineração de dados possibilitam o estudo de Mineração de Dados Adaptativa, que trata o processo de descoberta de padrões ocultos existentes em um determinado ambiente.

A proposta do modelo de Mineração de Dados Adaptativa é graficamente apresentada na figura 1.

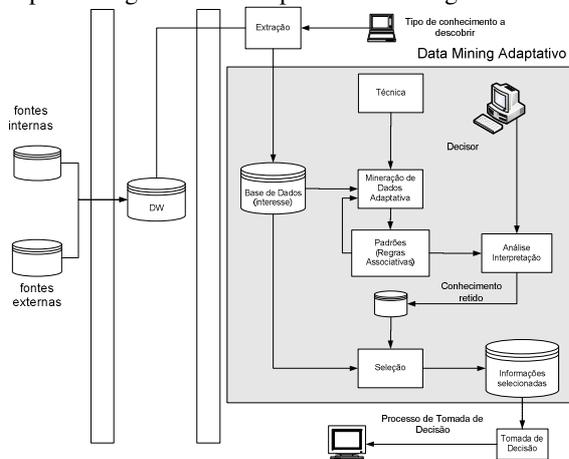


Figura 1 – Modelo de Mineração de Dados Adaptativa

O modelo proposto é composto pelas seguintes etapas principais:

1. definição do conhecimento a ser descoberto, que é fornecido por um decisor²;
2. a partir de uma base de dados organizada e da definição do padrão a ser descoberto, é feita a extração dos registros que satisfazem a definição solicitada;
3. é aplicada a mineração de dados para a descoberta de padrões, por exemplo, baseada em regras associativas, usando técnicas adaptativas; os resultados desta etapa resultam nos atributos associados, que formam o padrão descoberto;
4. a análise dos padrões obtidos, caso sejam validados pelo decisor, representam as regras associativas, que interpretadas são retidas como conhecimento adquirido;
5. o conhecimento retido, formado pelo conjunto de regras associativas, é usado para gerar uma base de dados com as informações relacionadas com a busca do conhecimento desejado;
6. a base de dados com as informações selecionadas pode ser utilizada, por exemplo, em um processo de tomada de decisão.

O modelo de Mineração de Dados Adaptativa (DMA) proposto é definido formalmente com o emprego da sêxtupla:

DMA = (BD, TC, TDM, PD, BC, FA), cujos elementos são descritos a seguir:

- BD é a base de dados composta pelos registros de interesse para a mineração de dados;
- TC é o tipo de conhecimento a ser descoberto;
- TDM é a técnica de mineração de dados a ser utilizada;
- PD é o conjunto de padrões descoberto;
- BC é a base de conhecimento resultante;
- FA representa o conjunto de funções adaptativas que forma a camada adaptativa do modelo.

A etapa de mineração de dados adaptativa consiste, portanto, na localização de padrões PD_i através de técnicas específicas TDM, como a de regras associativas, que através de funções adaptativas FA_j permitem gerar novas ocorrências de padrões. A aplicação das funções adaptativas quando instanciadas, executam as ações adaptativas de inserção, remoção e consulta no conjunto de regras, modificando de forma automática o próprio conjunto de regras.

O conhecimento em BC está codificado só no conjunto de regras. Uma transação adaptativa composta pelo conjunto de ações adaptativas leva a uma próxima situação, com uma configuração consistente, e incorpora uma nova informação no conjunto de regras.

²Neste estudo, decisor deve ser entendido como um especialista em determinado assunto, podendo ser um gestor, gerente, diretor ou até mesmo um grupo de profissionais.

Como exemplo de aplicação dos resultados obtidos no modelo proposto de Mineração de Dados Adaptativa, descreve-se nesta seção, de forma resumida, uma maneira de se integrar duas ferramentas adaptativas.

A figura 2 mostra como os resultados da Mineração de Dados Adaptativa podem servir a um processo de decisão com o auxílio de uma Tabela de Decisão Adaptativa para um determinado problema de tomada de decisão.

A Mineração de Dados Adaptativa apresenta ao final do seu processo um conjunto de regras associativas, que representam o conhecimento adquirido a ser utilizado em um processo de tomada de decisão.

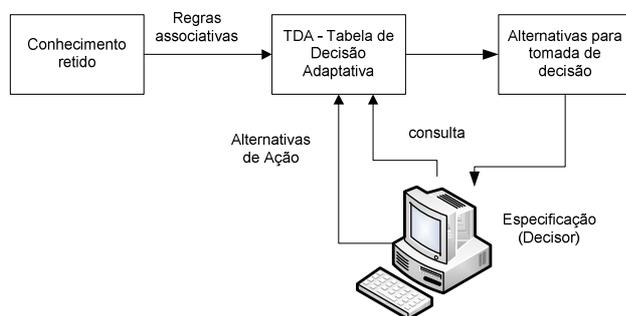


Figura 2 – Integração: Mineração de Dados Adaptativa e Tabela de Decisão Adaptativa

A tabela de decisão adaptativa, cujo formalismo se encontra em [6], pode apoiar nesse processo de tomada de decisão. A tabela de decisão pode ser gerada a partir das regras associativas, representando a entrada das regras iniciais R_k da tabela. Cada regra descreve o conhecimento descoberto na mineração de dados e, portanto, é composta pelo padrão encontrado. Esse padrão é composto por atributos, que representam os critérios C_i da tabela de decisão adaptativa.

Para completar os dados da tabela de decisão, o decisor deve inserir as alternativas de ação A_j para a tomada de decisão, associando-as às regras iniciais da tabela.

A tabela de decisão adaptativa, através de um método multicritério específico e da definição de funções adaptativas FAD_j relacionadas com as regras da tabela, está pronta para ser usada no processo de tomada de decisão.

Diante de um problema de decisão, consultas à tabela de decisão podem ser realizadas pelo decisor. A consulta tem como entrada uma regra RC , com atributos específicos, a ser pesquisada na tabela com o objetivo de buscar a solução ou alternativa de ação para a regra.

Uma ação adaptativa de consulta ao conjunto de regras da tabela é executada e caso os parâmetros da regra de entrada sejam coincidentes com alguma regra da tabela, então a alternativa de solução é apresentada.

Caso a regra não seja encontrada, uma ação adaptativa de inclusão de regra é executada, determinando qual é a alternativa de ação possível para a consulta. Uma nova configuração da tabela de decisão é apresentada, após

uma ação adaptativa de exclusão de regras que elimina regras conflitantes ou redundantes.

A nova configuração da tabela de decisão adaptativa resultante tem incorporado novas regras, podendo desta forma ser utilizada para novas consultas.

V. CONCLUSÕES

Neste artigo são apresentados os mecanismos de Mineração de Dados Adaptativa, originado em informações existentes em um determinado banco de dados, que contém registros de transações obtidos em fontes internas ou externas em um ambiente.

Esse dispositivo em estudo propõe um modelo genérico, podendo tratar diversas técnicas de mineração associadas às técnicas adaptativas, sendo sua aplicabilidade ampla, identificando padrões que geram conhecimento.

Para efeito de elaboração da proposta, esse estudo aborda o conceito Mineração de Dados e Tecnologia Adaptativa, que integrada a Tabela de Decisão Adaptativa, pode apoiar no processo de tomada de decisão.

Deve-se observar que essa proposta está em fase de estudo e de especificação.

O objetivo desta proposta é trazer contribuições, tanto nos aspectos teóricos e práticos de aplicação da Tecnologia Adaptativa quanto nos processos de Mineração de Dados na descoberta de padrões e tomada de decisão.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] NETO, J. J., *Adaptive Rule-Driven Devices - General Formulation and Case Study*. Lecture Notes in Computer Science. Watson, B.W. and Wood, D. (Eds.): Implementation and Application of Automata 6th International Conference, CIAA 2001, Vol. 2494, Pretoria, South Africa, July 23-25, Springer-Verlag, 2001, pp. 234-250.
- [2] RAMAKRISHNAN, R. *Sistema de Gerenciamento de Banco de Dados*. 3ª Ed. McGraw-Hill, São Paulo, 2008.
- [3] Watson, R. *Data Management: Banco de Dados e Organizações*. 3ª Ed. LTC, Rio de Janeiro 2004.
- [4] Kimbal, R., *The Data Warehouse Toolkit*. 2ª Ed. Wiley Computer Publishing, USA, 2002.
- [5] Moxon, B (1998). *Defining Data Mining-DBMS, Data Warehouse Supplement*, Aug.1996. [HTTP://dbmsmag.com/9608d53.html](http://dbmsmag.com/9608d53.html) (30.Out.2008).
- [6] TCHEMRA, A. H. *Aplicação da Tecnologia Adaptativa em Sistemas de Tomada de Decisão*. I WTA – Workshop sobre Tecnologia Adaptativa, Jan 2007 - www.pcs.usp.br/~lta.

Angela Hum Tchemra é bacharel e licenciada em Matemática pela Universidade Presbiteriana Mackenzie e mestre em Administração de Empresas pela mesma universidade. Atualmente é professora de cursos de graduação na Universidade Presbiteriana Mackenzie, Faculdade de Engenharia São Paulo e Fundação Armando Álvares Penteado. É Doutoranda do Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo.

Rubens de Camargo é bacharel em Administração de Empresas, com especialização em Análise de Sistemas pela Faculdade Associadas de São Paulo e mestre em Administração de Empresas pela Universidade Presbiteriana Mackenzie. Atualmente é professor de cursos de graduação na Universidade Presbiteriana Mackenzie. É Doutorando do Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo.