

# Usando Adaptatividade na Identificação de Padrões

(21 Janeiro 2010)

R. Camargo, Luís Raunheite

**Resumo.** Este artigo tem como objetivo, mostrar que para aplicações na área de reconhecimento de padrões, a adaptatividade pode em algumas situações, oferecer recursos que a tornam vantajosa em relação a outras técnicas. Entre as principais vantagens do uso da adaptatividade podem ser citadas: não impõe nenhuma notação particular, mas aceita qualquer formulação definida por regras; permite descrever de forma natural coisas dinâmicas complicadas; todos os dados úteis para o reconhecimento de padrões ficam disponíveis e restritos ao conjunto de regras; a evolução do conjunto de regras basta para um acompanhamento direto do aprendizado. Sua aplicabilidade será mostrada em um estudo a ser desenvolvido a partir de uma base de dados organizada, onde utilizando uma técnica de mineração de dados para elaboração de uma matriz de transição da cadeia de Markov e aplicando o conceito de convolução com procedimentos adaptativos, torna-se possível subsidiar processos de tomada de decisão.

**Palavras-chave:** Adaptatividade, padrões, conjunto de regras, base de dados, Markov.

## I. INTRODUÇÃO

A complexidade dos processos de identificação de padrões tem exigido estudos e pesquisas em diversas áreas, tanto na teoria como na prática em desenvolvimento de softwares.

Metodologias vêm sendo utilizadas para identificar padrões e obter conhecimento para posicionar empreendimentos de forma competitiva no mercado.

Este estudo aborda o uso dos formalismos adaptativos como uma alternativa para reconhecimento de padrões apresentado sua aplicabilidade onde haja necessidade, como por exemplo, auxiliar no processo de tomada de decisão.

Os sistemas computacionais que servem de apoio aos processos de tomada de decisão, geralmente, utilizam bancos de dados com informações coletadas previamente. Esses sistemas, além de apresentarem processos de realimentação de dados, permitem aos decisores criar vários cenários sobre um determinado problema de decisão [1]. Das análises dos

cenários e da decisão tomada, outras informações são obtidas e armazenadas para servirem de entrada para processos semelhantes.

De acordo com [2], pesquisas em inteligência artificial permitem o desenvolvimento de ferramentas computacionais inteligentes que auxiliam os usuários na execução de tarefas, entre elas, identificando padrões e auxiliando nos processos de tomada de decisão. Estas ferramentas são associadas à inteligência humana, como por exemplo, à capacidade de aprender, raciocinar e solucionar problemas.

Segundo [3], a Tecnologia Adaptativa é uma área da computação cujos estudos e pesquisas sobre técnicas adaptativas possibilitam aos dispositivos adaptativos apresentar como característica principal a capacidade de se automodificarem sem interferência externa, alterando suas estruturas topológicas, adaptando-se às necessidades requeridas de problemas específicos. Essa característica pode, desta maneira, conferir aos métodos adaptativos a classificação de sistemas inteligentes.

O objetivo deste artigo é apresentar o modelo no reconhecimento de padrões, baseado nos fundamentos da tecnologia adaptativa. Como parte dos estudos de pesquisa, é proposto um algoritmo para o modelo (desenvolvido em MatLab), que combina técnicas adaptativas ao processo de convolução.

## II. TECNOLOGIA ADAPTATIVA

Na área da Tecnologia Adaptativa existem inúmeros estudos de técnicas com aplicações em diversas áreas [3] [5], cujas contribuições têm estimulado o desenvolvimento de novos recursos computacionais.

Nos dispositivos adaptativos desenvolvidos, encontram-se formalismos conhecidos e tradicionais, tais como autômatos de pilha estruturados, *statecharts*, redes de Markov, gramáticas, árvores de decisão, tabelas de decisão, entre outros [5]. Isso mostra que há certa facilidade de uso das técnicas adaptativas, uma vez que [3] define um dispositivo adaptativo como um dispositivo formado por uma camada subjacente (núcleo do sistema) representada por um formalismo conhecido não-adaptativo e uma camada adaptativa, cujas funções agem sobre o núcleo, o que lhe confere a capacidade de automodificação.

As ações adaptativas são implementadas na camada adaptativa e são responsáveis pelas alterações no conjunto de

[1] R. Camargo – Universidade Presbiteriana Mackenzie (correspondência: R. Oscar Freire, 235 – ap.31- São Paulo, SP, Brasil – cep: 01426-001; e-mail: [rubens.camargo@poli.usp.br](mailto:rubens.camargo@poli.usp.br)).

[2] Luís Tadeu Mendes Raunheite – Universidade Presbiteriana Mackenzie; e-mail: [raunheite@mackenzie.br](mailto:raunheite@mackenzie.br)

regras, gerando uma nova configuração do dispositivo [3]. De maneira geral, as ações adaptativas permitem que regras sejam consultadas, eliminadas ou incluídas no sistema.

Um exemplo de dispositivo dirigido por regras adaptativo é a Árvore de Decisão Não-Determinística Adaptativa definida em [5]. O dispositivo adaptativo simula uma árvore de decisão, que é percorrida a partir da sua raiz (regra inicial) e chega a uma folha (caso determinístico) ou encontra vários ramos a serem percorridos (não-determinístico), após uma seqüência de testes em cada nó (regra). Quando não é possível atingir uma folha, ações adaptativas podem ser executadas na sub-árvore, realizando ações de consultas ou de remoção da sub-árvore ou de inclusão de uma nova sub-árvore à estrutura não-adaptativa.

O dispositivo dirigido por regras adaptativo definido em [3], tem como núcleo um processo que permite descobrir padrões com adaptatividade.

### III. CADEIA DE MARKOV

A rede de Markov pode ser vista como um sistema de estados e transições [8], semelhante a um autômato finito. Caracteriza-se de primeira ordem a probabilidade de um estado ser atingido dependendo apenas do estado atual, portanto um sistema estocástico, tratando a probabilidade entre estados, conforme Fig. 1.

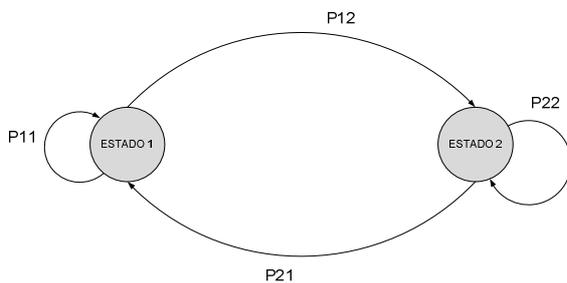


Fig. 1. transições de estados

Em [8], temos outra forma de representação do modelo de Markov a matriz de transição de estados, como sendo (fig. 2):

	Estado 1	Estado 2
Estado 1	0	1
Estado 2	0,49	0,51

Fig. 2. Quadro de transição de estados

Em outras palavras, a probabilidade do estado 1 seguir para o estado 2 é de 1 %, a probabilidade do estado 1 seguir para o estado 1 é de 0 %, a probabilidade do estado 2 seguir para o estado 1 é de 49 % e a probabilidade do estado 2 seguir para o estado 2 é de 51%.

Portanto:

$$\begin{matrix} p_{11} = 0 & p_{12} = 1 \\ p_{21} = 0,49 & p_{22} = 0,51 \end{matrix}$$

onde a probabilidade do estado 1 seguir para o estado 2 é de 49%, existindo uma probabilidade 51 %, ligeiramente maior do estado 2 seguir para o estado 2.

Conforme [7], a cadeia de Markov é um caso particular de processo estocástico, com tempo discreto, que segue a propriedade de Markov. A definição desta propriedade, também chamada de memória markoviana, é que os estados anteriores são irrelevantes para a predição dos estados seguintes, desde que o estado atual seja conhecido.

Uma cadeia de Markov é uma seqüência  $X_1, X_2, X_3, \dots$  de variáveis aleatórias. O escopo destas variáveis, isto é, o conjunto de valores que elas podem assumir, é chamado de espaço de estados, onde  $X_n$  denota o estado do processo no tempo  $n$ . Se a distribuição de probabilidade condicional de  $X_{n+1}$  nos estados passados é uma função apenas de  $X_n$ , então:

$$Pr(X_{n+1} = x | X_0, X_1, X_2, \dots, X_n) = Pr(X_{n+1} = x | X_n) \quad (1)$$

onde  $x$  é algum estado do processo. A identidade acima define a propriedade de Markov.

Seja  $P_{ij}$  a probabilidade condicional tal que se o sistema está no estado  $i$  em uma observação, então ele estará no estado  $j$  na próxima observação,  $1 \leq i \leq N, 1 \leq j \leq N$ . Essas probabilidades chamadas de probabilidades de transição. Para cada cadeia de Markov, a matriz  $P(N \times N)$  cujos elementos são  $P_{ij}$  é chamada de matriz de transição da cadeia de Markov (fig. 3), que no contexto da pesquisa será gerada a partir de valores obtidos no processo de mineração de dados.

$$P = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1N} \\ P_{21} & P_{22} & \dots & P_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N1} & P_{N2} & \dots & P_{NN} \end{bmatrix}$$

Fig. 3. Matriz de transição

No contexto da pesquisa o sistema em estudo é descrito através de uma matriz de transição da Cadeia de Markov, a partir das probabilidades de ocorrências de seus diferentes estados.

### IV. NÚCLEO DE CONVOLUÇÃO

Uma maneira de se entender a convolução é como uma operação que copia uma matriz a partir de cada localização de elemento para outra, considerando o valor de todos os elementos na área onde a cópia acontece, para produzir a alteração no valor original do elemento da matriz, [4].

Pode-se descrever a convolução como um processo de somas ponderadas. Cada elemento da matriz na vizinhança é multiplicado pelo seu similar no núcleo de convolução; a soma de todos os produtos resulta no novo valor do elemento central de interesse. Cada elemento do núcleo de convolução é um fator de ponderação (também chamado de coeficiente de convolução). O arranjo dos fatores de ponderação no núcleo, bem como o tamanho do núcleo, determina o tipo de transformação que será aplicada ao dado da representação matricial. Mudando um fator de ponderação no núcleo de convolução muda-se a magnitude e até o sinal de toda a soma afetando o valor atribuído ao elemento de interesse.

A convolução por soma ponderada apresenta um problema na sua aplicação nas fronteiras da matriz. Com o movimento do núcleo de convolução através da matriz, ao chegar a fronteira da mesma quando o elemento de interesse estiver na fronteira, uma parte dos coeficientes do núcleo não estarão sobre elementos da matriz. Uma maneira de contornar este problema é ignorar as fronteiras da matriz no cálculo, outra é duplicar os dados da fronteira, de forma a adicionar uma fronteira à matriz original, permitindo assim o cálculo sobre a fronteira original.

A operação de convolução substitui o valor do elemento da matriz pela soma de seu valor com a valor dos elementos das vizinhanças, tudo multiplicado por um fator chamado de "núcleo de convolução", supondo que se use uma vizinhança de 3X3 elementos, chamando de  $p(x,y)$  os pontos de matriz e os pontos do núcleo de  $N(x,y)$  onde  $x = 0, 1$  ou  $2$ , então o elemento central,  $p(1,1)$  será substituído pela soma dos pontos, vezes o valor do núcleo.

$$p(1,1) = p(0,0) * N(0,0) + p(1,0) * N(1,0) + p(2,0) * N(2,0) + p(0,1) * N(0,1) + p(1,1) * N(1,1) + p(2,1) * N(2,1) + p(0,2) * N(0,2) + p(1,2) * N(1,2) + p(2,2) * N(2,2) \quad (2)$$

ou

$$p(1,1) = \sum_{m,n} N(m,n) * p(m,n) \quad (3)$$

Esta é uma operação de correlação. Para convolução pode-se inverter a ordem dos valores do núcleo. A correlação é mais fácil de se entender e muitas convoluções de núcleo são simétricas, sendo equivalentes à correlação. Para convoluir uma área da matriz, deve-se repetir esta operação para cada posição de elemento na matriz de dados. Em cada ponto, deve-se multiplicar os valores do núcleo com os valores da matriz sobre ela, somar o resultado, e substituir o elemento do centro do núcleo com o valor. A equação então se torna:

$$p(x,y) = \sum_{m,n=0}^2 N(m,n) * p(x+m,y+n) \quad (4)$$

Convoluir uma área de tamanho X por Y com um núcleo de tamanho m por n requer  $X*Y*m*n$  multiplicações e somas.

Então uma matriz de 256 por 256 com um núcleo de 3 por 3 requer 589.824 multiplicações e somas. A Fig. 4 representa o processo de convolução.

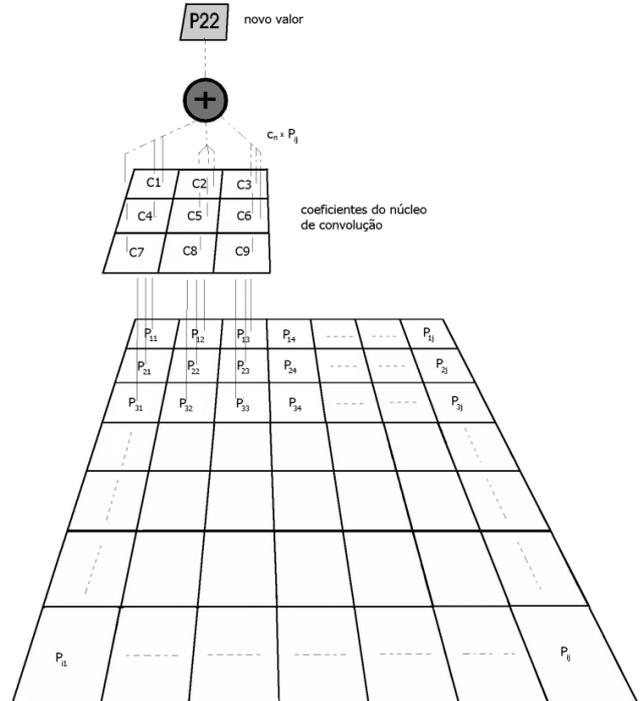


Fig. 4. Aplicação do processo de convolução

Quando se aplica a convolução para um problema de processamento de dados pensa-se na convolução como um filtro espacial. Em um filtro espacial, a convolução de núcleo é essencialmente para análise de uma pequena parte da matriz de dados que se quer amplificar ou detectar. Nesse contexto a escolha eficiente do núcleo pode detectar características na matriz de dados que permitam reconhecer padrões a partir da diferença de frequência dos resultados obtidos. Através da identificação de um determinado padrão o método proposto permitirá alterações no núcleo de convolução o que torna adaptativa a técnica. Essa identificação será feita a partir de um algoritmo que tomará como endereço de uma tabela de consulta podendo ser uma tabela de um banco de dados, este será o índice obtido na convolução, portanto este conteúdo do endereço indicado na tabela conterà os novos valores a serem usados como novo núcleo de convolução adaptativo

## V. CONCEITOS DE MINERAÇÃO DE DADOS

Abordando pesquisas que tem como objetivo básico a identificação de padrões encontram-se estudos sobre Mineração de Dados ou Data Mining que 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 [9].

Mineração de dados, segundo [10], 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 [11], 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 [11], é uma unidade de informação ou atributo de um registro que se repete, ou então é uma sequência de informações/atributos presentes em uma estrutura que se repete.

Existem cinco tipos de técnicas [10], 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 Sequenciais: que identificam sequências 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 frequência do comparecimento de clientes em uma agência bancária: infrequentes (nunca frequentam a agência), frequentes (comparecem de modo frequente) e ocasionais (ocasionalmente frequentam 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.

considerando a “experiência anterior” adquirida por um dispositivo adaptativo, baseado em um histórico de operações, um sistema pode tomar a decisão de modificar seu comportamento sem a inferência de qualquer agente externo. Conforme [5], o formalismo adaptativo se mostra uma opção a ser utilizada na aprendizagem computacional, destacando trabalhos realizados no reconhecimento de imagens, linguagens. Tratando o reconhecimento de padrões, [5] ainda destaca o reconhecimento ótico de caracteres (OCR), baseado em classes de técnicas de aprendizagem tais como: árvores de decisão, redes neurais artificiais, sentenças em lógica de predicados, conjunto de regras “se-então”, autômatos, redes bayesianas e memorização (*instance-based learning*). Com relação as regras “se-então”, pode-se entender como uma opção de fácil entendimento sem o domínio de conceitos complexos, permitindo uma maior transparência do modelo.

Uma proposta de trabalho com a utilização do conceito de adaptatividade no reconhecimento de padrões em uma base de dados, tem como abrangência os seguintes elementos:

- a) a utilização da técnica de associação usada na Mineração de dados;
- b) a partir de interações que classificam as informações armazenadas, será definida uma camada de aplicação da rede de Markov na identificação das decisões de agrupamento dos dados;
- c) definição de uma camada de Convolução que aplicada sobre a rede de Markov gera o fator de ponderação, que determina o tipo de transformação aplicada na representação matricial.

Esta última camada será denominada Convolução Adaptativa, pois apresentará seus fatores de ponderação de forma variável, havendo a possibilidade de inclusão ou exclusão de novos fatores de ponderação, alterando os coeficientes do núcleo de convolução, face aos resultados obtidos (fig. 5).

## VI. RECONHECENDO PADRÕES COM ADAPTATIVIDADE

Adaptatividade conforme [6], refere-se a um conceito onde

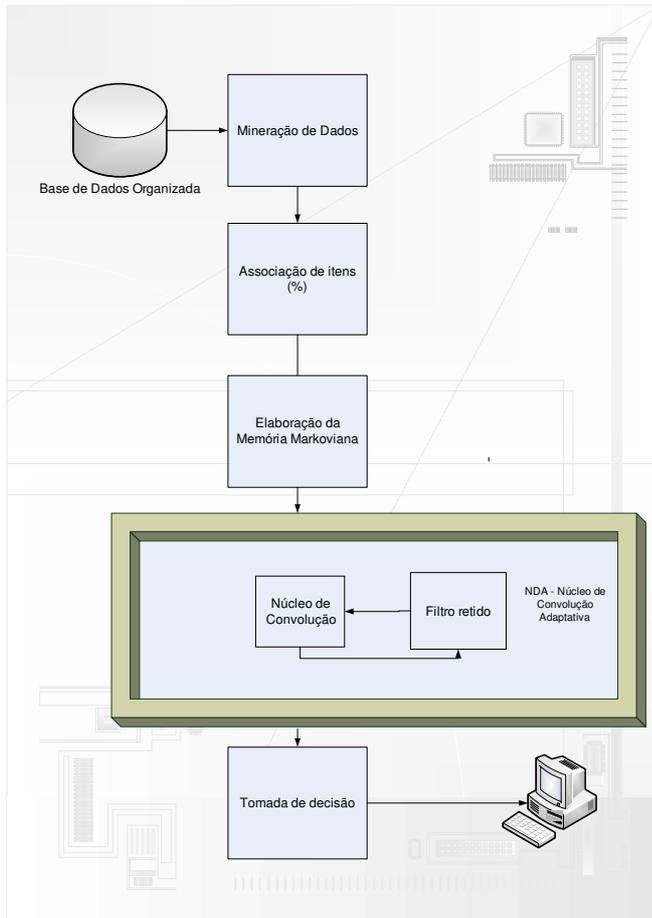


Fig. 5. Modelo do Estudo

VII. MODELO DE APLICAÇÃO

Como modelo hipotético de aplicação da técnica adaptativa proposta foi utilizada uma Matriz de Transição (5), em duas transições de estados, e calculada a convolução considerando o núcleo (6).

Matriz de Transição da Cadeia de Markov

$$\begin{bmatrix} 0.3750 & 0.5000 & 0.1250 & 0.0000 \\ 0.250 & 0.4375 & 0.2500 & 0.0625 \\ 0.0625 & 0.2500 & 0.4375 & 0.2500 \\ 0.0000 & 0.1250 & 0.5000 & 0.3750 \end{bmatrix} \quad (5)$$

Núcleo hipotético de Convolução:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \quad (6)$$

Como resultado da convolução temos:

$$\begin{bmatrix} 2.6875 & 3.0000 & 2.3125 & 1.1250 \\ 3.3125 & 4.0000 & 3.3125 & 2.0000 \\ 2.3125 & 3.0000 & 2.6875 & 1.8750 \\ 1.3750 & 2.0000 & 1.9375 & 1.5625 \end{bmatrix} \quad (7)$$

Calculando o histograma para o resultado obtido:

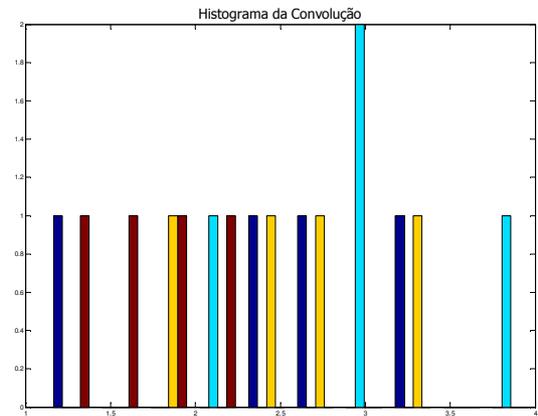


Fig. 6. Histograma da convolução

Pode-se notar que o histograma da convolução (Fig. 6), aplicada à Matriz de transição destaca uma determinada frequência espacial que pode ser usada como referência em nova aplicação da técnica alterando, por exemplo, o núcleo de convolução.

O resultado da convolução será comparado com níveis pré-estabelecidos que indicarão a necessidade da utilização de novos núcleos de convolução, caso afirmativo um ponteiro será definido usando esse resultado. Assim o ponteiro indicará o endereço onde os coeficientes do novo núcleo estarão armazenados.

VIII. VII. CONCLUSÃO

Neste artigo são apresentados mecanismos adaptativos de identificação de padrões, objetivando a criação de um aplicativo de apoio a tomada de decisão, utilizando conceitos de mineração de dados, memória Markoviana, convolução e adaptatividade. É objetivo desta proposta criar um dispositivo adaptativo para aplicação em caráter genérico em processos de tomada de tomada de decisão.

REFERÊNCIAS BIBLIOGRÁFICAS

[1] PIDD, M., *Modelagem empresarial: ferramentas para tomada de decisão*. Trad. Gustavo Severo de Borba et al. Porto Alegre: Artes Médicas, 1998.  
 [2] O'BRIEN, J. A. *Sistemas de Informação e as Decisões Gerenciais na Era da Internet*. 2ed. Saraiva, São Paulo, 2004.  
 [3] 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.
- [4] HU, Osvaldo; RAUNHEITTE, Luís, *Processamento e Compressão Digital de Imagens*. Editora Mackenzie, São Paulo, 2004.
- [5] PISTORI, H. *Tecnología Adaptativa em Engenharia de Computação: Estado da Arte e Aplicações*. Tese de Doutorado – Escola Politécnica da USP, 2003.
- [6] NETO, J.J., *Um Levantamento da Evolução da Adaptatividade e da Tecnologia Adaptativa*. IEEE Latin America Transactions, Vol. 5, No. 7, Nov. 2007.
- [7] MAKI, Daniel; THOMPSON, Maynard, *Mathematical Modeling and Computer Simulation*, Thomson Brooks/Cole, USA, 2006.
- [8] BASSETO, Bruno Arantes., *Um Sistema de Composição Musical Automatizada, Baseado em Gramáticas Sensíveis ao Contexto, Implementado com Formalismos Adaptativos*. Dissertação de Mestrado - Escola Politécnica da USP, 2000.
- [9] Kimbal, R., *The Data Warehouse Toolkit*. 2ª Ed. Wiley Computer Publishing, USA, 2002.
- [10] 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).
- [11] RAMAKRISHNAN, R. *Sistema de Gerenciamento de Banco de Dados*. 3ª Ed. McGraw-Hill, São Paulo, 2008.

**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 Faculdade de Computação e Informática da Universidade Presbiteriana Mackenzie. É Doutorando do Departamento de Engenharia de Computação e Sistemas Digitais da Escola Politécnica da Universidade de São Paulo.

**Luís Tadeu Mendes Raunheitte** é Engenheiro Eletrônico pela Universidade Presbiteriana Mackenzie, mestre e doutor em Engenharia Elétrica pela mesma Universidade. Atualmente é professor de cursos de graduação na Faculdade de Computação e Informática da Universidade Presbiteriana Mackenzie.