



## XVIII Seminário Nacional de Distribuição de Energia Elétrica

SENDI 2008 - 06 a 10 de outubro

Olinda - Pernambuco - Brasil

### **Inteligência Computacional nas Distribuidoras de Energia Elétrica: evolução tecnológica, aplicações e impactos na redução das perdas não-técnicas**

<b>Bruno L. A. Alberto</b>	<b>Paulo E. Maciel de Almeida</b>	<b>Rodrigo Leite Durães</b>
<b>CEFET-MG</b>	<b>CEFET-MG</b>	<b>CEFET-MG</b>
brunol@lsi.cefetmg.com.br	pema@lsi.cefetmg.com.br	rodrigo@lsi.cefetmg.com.br

#### **Palavras-chave**

Inteligência Computacional  
Mineração de Dados  
Redes Neurais Artificiais  
Inteligência Artificial  
Perdas não-técnicas

#### **Resumo**

Este trabalho apresenta as principais aplicações de Inteligência Computacional e seus impactos na redução de perdas não-técnicas nas Distribuidoras de Energia Elétrica no país. São apresentadas as últimas pesquisas e projetos realizados na área, bem como os últimos trabalhos publicados em congressos e seminários nacionais e internacionais ocorridos no país, tais como Metering Latin America, Sendi e Citenel. A Inteligência Computacional é um campo do conhecimento que oferece modelos de apoio à decisão com base na descoberta de informações relevantes e presentes, embora muitas vezes não aparentes, em grandes quantidades de dados armazenados em repositórios de informação e bancos de dados. Baseada em fatos reais, conhecimentos empíricos e teóricos, a Inteligência Computacional vem obtendo grande interesse por parte das Distribuidoras, grandes consumidores e fornecedores de soluções tecnológicas para o mercado de energia.

#### **1. Introdução**

O termo Inteligência Computacional (IC) pode ser generalizado como a simulação da “inteligência” humana para realização de tarefas até então elaboradas e executadas por pessoas e que podem ser resolvidas pelo uso de computadores e algoritmos inteligentes.

A filosofia da IC está ligada ao estudo das áreas do conhecimento que procuram compreender o processo de como fazer os computadores realizarem tarefas, que no momento, as pessoas fazem melhor. Inicialmente, ela estava focada em tarefas formais, tais como jogos, demonstrações de teoremas e raciocínios do senso comum. Posteriormente, foi utilizada com uma maior manipulação de conhecimentos, com algoritmos de percepção (visão e fala), compreensão da linguagem natural e a solução de problemas em domínios especializados, como diagnósticos médicos, análise química e

detecção de padrões. Atualmente, a IC está sendo também utilizada em projetos na área de engenharia, descobertas científicas, planejamento financeiro, mineração e análise de dados para a geração de informações empresariais, sejam em Sistemas Especialistas ou Sistemas de Informação.

O objetivo no desenvolvimento contemporâneo de sistemas de IC não é substituir completamente a tomada de decisões humana, e sim reaplicá-la em certos tipos de problemas bem definidos. Assim como em outros Sistemas de Informação, o propósito maior das aplicações da IC nas empresas é auxiliar as mesmas a alcançarem suas metas. A IC envolve o conceito de explorar um comportamento inteligente que possui características definidas como aprender com a experiência, aplicar o conhecimento adquirido da experiência, tratar situações complexas, resolver problemas quando faltam informações importantes, determinar o que é importante, ter capacidade de raciocinar e pensar, reagir rápida e corretamente a novas situações, compreender imagens visuais, processar e manipular símbolos, ser criativo e imaginativo e utilizar da heurística (normas práticas advindas da experiência). Este vasto campo possui diversos componentes importantes, como sistemas especialistas, robótica, sistemas de visão de imagens, processamento de linguagem natural, sistemas de aprendizagem, lógica difusa, redes neurais, algoritmos genéticos e agentes inteligentes.

Este artigo está dividido em cinco seções. A seção dois apresenta os principais projetos, trabalhos e pesquisas na área de IC aplicada no gerenciamento das perdas não-técnicas nas Distribuidoras de Energia nos últimos anos no país. Na seção três são apresentados os benefícios gerados pela aplicação de IC no gerenciamento das perdas não-técnicas nessas empresas e as conclusões deste trabalho. As referências bibliográficas são apresentadas na seção quatro.

## **2. Inteligência Computacional no gerenciamento das perdas não-técnicas**

Gerenciar conhecimento de especialistas em perdas não-técnicas é um processo extremamente importante para as Distribuidoras de Energia. Sob um cenário dinâmico e com taxas tendenciosamente crescentes de perdas, o controle manual, típico dos processos de gestão de perdas não-técnicas atuais, é exercido instante a instante por intermédio de decisões que exigem do especialista o conhecimento necessário para relacionar ações e resultados. É exigido também deste profissional o acompanhamento da operação, análise de grande base de dados, relacionamento com ambientes e áreas diferentes, experiência e segurança suficiente para assumir riscos em situações extremas. Há ainda o dilema da assertividade: manter o número de inspeções e aumentar a assertividade ou reduzir a necessidade de inspeções e manter a assertividade.

Um dos grandes desafios da gestão das perdas não-técnicas está relacionado ao volume de dados disponível aos especialistas para a execução das análises e tomada de decisão. O número de aplicações, sistemas e repositórios de informações que coexistem dentro das Distribuidoras tem crescido rapidamente e muitas vezes estes são desenvolvidos para atender aos requisitos específicos de cada um dos setores da empresa, sejam eles comerciais (faturamento), relacionamento com clientes ou mesmo os sistemas eletrônicos de medição. Apesar do seu crescimento e grande aceitação, tais sistemas e repositórios ainda apresentam baixa reutilização de componentes, falta de melhores práticas de manutenção, problemas de interoperabilidade e tempo de desenvolvimento demasiadamente alto. Além disso, existe fragmentação dos dados, dificuldade de obtenção de informações confiáveis e consolidadas, redundância e inconsistência de dados armazenados em mais de um sistema. Outro problema igualmente relevante é o fato de que utilizando sistemas fragmentados, os usuários criam barreiras para o fluxo das informações dentro da empresa, pois em particular se enxergam como detentores do conhecimento inerente ao seu setor ou à função que desempenham.

É neste cenário que a IC se insere como um conjunto de técnicas a serem aplicadas na extração de conhecimento e tomada de decisão sobre grandes bases de dados. A IC é um campo do conhecimento que oferece modelos de apoio à decisão e ao controle com base em fatos reais e conhecimentos empíricos e teóricos, mesmo que apoiados em dados incompletos. Na literatura nacional, são verificados diversos projetos, trabalhos e pesquisas realizados nos últimos anos nesta área, os quais apresentam resultados interessantes para as Distribuidoras de Energia.

Em Francisco et al. (2006, p. 1-8), é verificada a aplicação e análise da metodologia *KDD (Knowledge Discovery in Databases)* por meio de um software aplicativo da empresa SPSS™ para segmentação de clientes de baixa tensão da AES Eletropaulo. A partir de dados técnicos e comerciais advindos dos diversos processos de gestão da empresa e de fontes externas, entre elas Censo Demográfico 2000 IBGE, pesquisa ABRADDEE 2005 do cliente residencial urbano e localização geográfica dos clientes da AES Eletropaulo, é possível identificar segmentos de clientes para tomadas de decisão estratégica de marketing, redução de inadimplência e fraude. Neste trabalho, foram verificadas e analisadas as técnicas *Two Step Cluster* e Critério de Informação de *Bayes* para a criação de 14 clusters, entre eles o de “fraudador“. Este informa à empresa um perfil de cliente que realiza chamadas ao *Call Center* para serviços e informações, que apresenta serviço não regulado e que possui débito automático. Neste caso, é recomendado um tratamento diferenciado nos canais de relacionamentos, aliado à antecipação de ações fraudulentas e contínuo monitoramento dos pagamentos com restrição de concessão aos serviços não regulados e débito automático.

Em Todesco et al. (2007, p. 1-8), é apresentado um sistema para identificação de possíveis fraudadores de energia elétrica empregando também a técnica de *KDD*. Foram utilizadas somente informações sobre o consumo para calcular a diferença entre o consumo do mês atual e o consumo do mesmo mês no ano anterior para 12 meses, sendo acumulado no último mês. Após ajuste do sistema, a taxa de acerto para o grupo de consumidores residenciais foi de 64% e a taxa média de acerto para o grupo de consumidores comerciais (padarias, lanchonetes e postos de gasolina) foi de 80%.

Moraes & Silveira (2006, p. 1-8) apresentam o processo de gestão de perdas não-técnicas da distribuidora de energia Elektro, a qual desenvolveu um sistema baseado no histórico de consumo, em informações cadastrais e na utilização de filtros para auxiliar no direcionamento da geração das ordens de serviço. Juntamente com os procedimentos gerenciais adotados pela empresa, o sistema permitiu à empresa reduzir suas perdas não-técnicas e manter o índice de acerto de irregularidades em 12%. Ainda em fase de desenvolvimento, estão sendo testados modelos de redes neurais e árvores de decisão com intuito de aumentar a assertividade do sistema.

Segundo Prati et al. (2003, p. 1-7), uma vez que o número de não-fraudadores geralmente é menor que o número de fraudadores, o treinamento de redes neurais deve ser considerado tendo em vista o desbalanceamento das classes de dados. Quando a diferença entre o número de exemplos pertencentes a cada uma das classes é grande, podem ocorrer dificuldades em induzir o conceito relacionado à classe minoritária. Nessas condições, sistemas de classificação que são otimizados em relação à precisão têm tendência de criar modelos triviais, que quase sempre predizem a classe majoritária.

Passini & Toledo (2007, p. 1-8) propõem o uso de mineração de dados para identificar fraudes em ligação de água na cidade de Campinas, mais especificamente na empresa de saneamento Sanasa. Embora aqui não se discuta uma aplicação no setor elétrico, este trabalho, gerado a partir da dissertação de mestrado de Passini, propõe soluções perfeitamente customizáveis ao setor elétrico. Baseado em dados históricos armazenados há mais de dez anos em banco de dados, foi utilizado o

software DB2 Intelligent Miner© (IM) da empresa IBM™ e aplicada as diversas técnicas que o compõem para levantamento do perfil típico de um fraudador e o tipo de fraude executada. São técnicas utilizadas pelo DB2 IM: associação, classificação por árvore de decisão, classificação neural, agrupamento demográfico, agrupamento neural, predição *RBF* (*Radial-Basis-Function*), predição neural, padrões sequenciais e seqüências similares. Visando inicialmente uma redução de 51% para 41% das inspeções mal-sucedidas, foram aplicadas as técnicas de redes neurais e árvores de decisão, as quais não alcançaram os resultados planejados. Atribuíram-se, entre outros, como fatores responsáveis pelo alto índice de erro, a escolha das variáveis de entrada do processo e a falta de testes utilizando outros modelos e técnicas presentes no software.

Gemignani et. al (2006, p. 1-8) apresenta as etapas de desenvolvimento de um projeto de P&D (Pesquisa e Desenvolvimento) executado pela distribuidora de energia Eletropaulo, o qual consiste na elaboração de uma metodologia e respectiva implementação computacional para a estimativa de consumo de clientes fraudadores de energia elétrica através da determinação dos fatores de carga e de demanda típicos para cada atividade e classe de consumo. A partir da potência instalada do cliente fraudador, obtida através de inspeção da carga no ato de constatação da fraude, e dos fatores de carga e de demanda definidos pela metodologia desenvolvida, pode-se estimar a energia consumida não faturada dos clientes com instalações irregulares. Foi utilizada a técnica de regressão linear pelo método de mínimos quadrados.

Patrício (2005, p. 10-86) em sua dissertação de mestrado apresenta uma metodologia que define perfis de comportamentos diários de unidades consumidoras de energia elétrica ligadas em alta tensão, com a finalidade de detectar fraudes ou erros de medição. A partir desta metodologia, construiu-se um sistema baseado em regras, utilizando informações estáticas (dados cadastrais) e dinâmicas (memória de massa dos medidores obtida de forma on-line através de sistemas de telemedição instalados nos clientes) sobre o consumidor. Para o desenvolvimento deste sistema foi utilizada a teoria de *KDD* para mineração de dados inicial e teoria de *Rought Sets* para seleção de atributos relevantes e geração de regras. Classificou-se o comportamento das unidades, através dos perfis diários ou semanais, em normais e anormais. Os clientes classificados como anormais são selecionados para inspeção técnica. Os resultados foram considerados satisfatórios, uma vez que a taxa de acerto na identificação de fraude obtida pelo sistema, utilizando-se análise semanal na unidade consumidora, a partir da pré-seleção dos consumidores com suspeitas de fraude foi de 64,7%.

Em Oltremari et. al (2006, p. 1-8), são apresentados dois algoritmos de classificação aplicados na detecção de consumidores fraudadores de energia: o classificador *Naïve Bayesian* e árvore de decisão (*decision tree*). Estes consistem na extração de padrões de consumo em kWh de dados históricos e arranjo dos dados de várias formas, calculando a média anual, mensal, semanal e diária. O objetivo consiste em assegurar resultados ótimos no desenvolvimento de perfis de carga para serem usados como referência na análise de perdas não-técnicas. Foram testadas duas arquiteturas de Rede Neural do tipo *Multi-Layer Perceptron* com algoritmo de treinamento *Backpropagation* e utilização de cinco variáveis de entrada: Demanda Média Medida, Desvio Padrão da Demanda Medida, Consumo Médio, Demanda Contratada Média e Fator de Carga Médio para a primeira Rede (RNA-1) e quatro variáveis de entrada: Demanda Medida, Demanda Contratada, Fator de Carga e Potência Instalada para a segunda Rede (RNA-2). Ficou estabelecido que a saída de cada rede fosse uma variável do tipo binária, indicando se o consumidor analisado é suspeito de fraude ou não. Com uma pequena massa de dados para treinamento, teste e validação, obteve-se para as duas arquiteturas o erro de classificação em torno de 30%.

No trabalho de Eller (2003, p. 10-92), o problema da fraude na área de distribuição de energia também é tratado aplicando-se uma rede neural do tipo *Multi-Layer Perceptron* com algoritmo de treinamento *Backpropagation* para classificação de consumidores fraudulentos residenciais/comerciais e segmentação com redes do tipo *Self Organizing Maps* para consumidores industriais. São utilizadas bases de dados reais da distribuidora de energia Celesc, constituindo-se de informações relativas ao histórico de consumo de seus consumidores residenciais, comerciais e industriais, além de variáveis como tipo do medidor, demanda e consumo por período. Após a análise, o resultado das tarefas de classificação e segmentação de consumidores fraudadores alimenta aplicativos responsáveis por disparar ordens de inspeção para averiguar a ocorrência da fraude. Em um dos experimentos para classificação de consumidores comerciais fraudadores, de 24 ordens de inspeção disparadas pela análise do software de redes neurais, apenas 2 se confirmaram como fraude após o retorno das equipes de inspeção de campo.

#### **4. Considerações finais**

Este trabalho apresentou as principais aplicações de Inteligência Computacional e seus impactos na redução de perdas não-técnicas nas Distribuidoras de Energia Elétrica no país durante os últimos anos. Foram apresentadas as pesquisas e projetos mais recentes realizados na área, bem como os últimos trabalhos publicados em congressos e seminários nacionais e internacionais ocorridos no país, tais como Metering Latin America, Sendi, e Citenel.

Embora a Inteligência Computacional tenha se mostrado um campo vasto para pesquisa aplicada, ainda é pouco utilizada nas Distribuidoras de Energia no que diz respeito a auxiliar os analistas no gerenciamento das perdas não-técnicas. Os sistemas de informação e repositórios de dados ainda apresentam baixa reutilização de componentes, problemas de interoperabilidade e manutenção, dificuldade de obtenção de informações confiáveis e consolidadas (principalmente no cadastro de consumidores) e redundância e inconsistência de dados armazenados em mais de um sistema. Além disso, existem barreiras criadas pelos próprios usuários para o fluxo das informações dentro da empresa, pois em particular se enxergam como detentores do conhecimento inerente ao seu setor ou à função que desempenham.

Das técnicas mais utilizadas e promissoras no gerenciamento das perdas não-técnicas utilizando IC nas Distribuidoras de Energia, estão as Redes Neurais, o *KDD* e as árvores de decisão. Embora ainda apresentem resultados modestos, ambas têm se mostrado mais populares neste tipo de atividade nas empresas. Na maior parte das pesquisas, informações de consumo (histórico) e dados estáticos provenientes dos sistemas comerciais das empresas (cadastro e faturamento) são tidos como as principais variáveis de entrada dos projetos e trabalhos verificados.

Uma das dificuldades encontradas é a falta de consistência principalmente nos dados de cadastros dos consumidores em análise. Outra é a falta de variáveis temporais em intervalos menores, principalmente aquelas provenientes dos consumidores atendidos em baixa tensão de distribuição, os quais têm suas informações de consumo coletadas apenas uma única vez ao mês e em intervalos relativos a um mês de consumo. Para os clientes atendidos em média e alta tensão, é grande o número de sistemas de medição eletrônica, os quais provêm, além de um maior número de informações sobre o consumidor em questão (consumo, potência, tensão, corrente, fator de potencia), leituras em intervalos de tempo bem menores, na casa de minutos, proporcionadas por sistemas de comunicação via rádio frequência.

## 5. Referências bibliográficas

ALMEIDA, P.E.M, MEIRELES, M. R. G;. & SIMÕES, M.G. A Comprehensive Review about Industrial Applicability of Artificial Neural Networks. USA, IEEE Transactions on Industrial Electronics, V.50, N.3, 2003, p.585-601.

PAULINO, Clóvis A. Estudo de tecnologias aplicáveis à automação da medição de energia elétrica residencial visando minimização de perdas. São Paulo, USP, Dissertação de Mestrado, 2006, p. 19-89

VIEIRALVES, Eduardo de Xerez. Proposta de uma metodologia para avaliação das perdas comerciais dos sistemas elétricos: o caso Manaus. São Paulo, Unicamp, Dissertação de Mestrado, 2005, p. 2-122.

FRANCISCO, Eduardo de Rezende, PETRIELLI, Adriana & REINA, Cláudia Sciortino. Segmentação comportamental de clientes para o setor elétrico. São Paulo, In: Anais do III Congresso anual de tecnologias da informação da Fundação Getúlio Vargas, 2006, p. 1-8

TODESCO, J. L., GARBELOTO, L. A., MORALES, A. B. T., ATHAYDE, E. D., RAUTENBERG, S. Aplicação de técnicas de Mineração de Dados para detecção de fraudes de energia. Araxá, In: Anais do IV CITENEL Congresso de Inovação Tecnológica em Engenharia Elétrica, 2007, p. 1-8.

PRATI, R. C., BATISTA, G. E. A. P. A &, MONARD, M. C. Uma experiência no balanceamento artificial de conjuntos de dados para aprendizado com classes desbalanceadas utilizando análise ROC. Chile, In: Anais da IV Jornada Chilena de Computação, IV Workshop de Inteligência Artificial, 2003, p. 1-10.

MORAES, Rodrigo Teodoro Bilia & SILVEIRA, Alex. Otimizando Resultados com a Gestão das Perdas Não-Técnicas de Energia Elétrica. Belo Horizonte, In: Anais do XVII SENDI Seminário Nacional de Energia Elétrica, 2006, p. 1-8.

PASSINI, Sílvia Reginato & TOLEDO, Carlos Miguel Tobar. Rodrigo Teodoro Bilia & SILVEIRA, Alex. Mineração de Dados para Detecção de Fraudes em Ligações de Água. Blumenau, In: XVI Seminário de Computação, 2007, p. 1-8.

GEMIGNANI, M. M. F. et al. Desenvolvimento de Sistema de Estimativa de Consumo para Recuperação de Receitas. Belo Horizonte, In: Anais do XVII SENDI Seminário Nacional de Energia Elétrica, 2006, p. 1-8.

PATRÍCIO, M. M. M. Detecção de fraude ou erro de medição em grandes consumidores de energia elétrica utilizando Rough Sets baseado em dados históricos e em dados em tempo real. Campo Grande, Universidade Federal de Mato Grosso do Sul, Dissertação de mestrado, 2005.

OLTREMARI, A. M. et al. Desenvolvimento de um Sistema Especialista para Detecção de Pontos Potenciais de Perdas Comerciais. Belo Horizonte, In: Anais do XVII SENDI Seminário Nacional de Energia Elétrica, 2006, p. 1-8.

ELLER, N.A. Arquitetura de informação para o gerenciamento de perdas comerciais de energia elétrica. Florianópolis, Tese de Doutorado, Universidade Federal de Santa Catarina, Departamento de Engenharia de Produção e Sistemas, 2003, p. 10-95.