

Serviços solicitados por consumidores atendidos em baixa tensão junto às distribuidoras de energia elétrica: formação de clusters com suspeita de irregularidades utilizando Mapas Auto-Organizáveis

Bruno Lambertucci Araújo Alberto

Departamento de Pesquisa e Desenvolvimento, DPD
SENERGY Sistemas de Medição S/A
30.170-020, Belo Horizonte, MG
E-mail: bruno@senergy.com.br

Paulo Eduardo Maciel de Almeida

Laboratório de Sistemas Inteligentes, LSI
Centro Federal de Educação Tecnológica de Minas Gerais, CEFET-MG
30.510-000, Belo Horizonte, MG
E-mail: pema@lsi.cefetmg.br

RESUMO

Este artigo descreve a aplicação dos Mapas Auto-Organizáveis para formação de agrupamentos de unidades consumidoras atendidas em baixa tensão de distribuição com suspeita de irregularidades no consumo de energia elétrica. Foi implementado um estudo de caso envolvendo uma cidade brasileira com aproximadamente 90 mil consumidores no período entre os anos de 2003 e 2007 cujos resultados encontrados foram promissores para potencialização de esquemas de classificação de padrões de comportamento fraudulentos de consumidores de energia.

Palavras-chave: *Mapas Auto-Organizáveis, Classificação de Padrões, Perdas Comerciais.*

Introdução

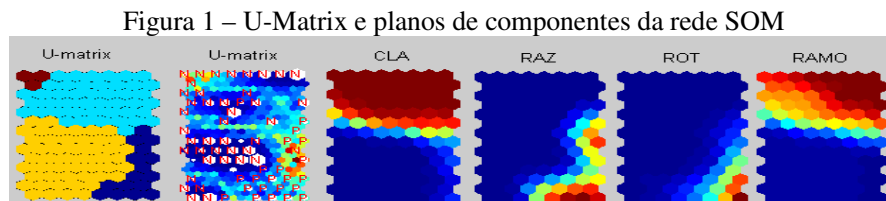
Os Mapas Auto-Organizáveis (SOM) têm sido aplicados, com sucesso, em variados problemas de análise de dados multidimensionais [1] [2], todavia, poucos são os trabalhos voltados para a análise multidimensional de informações coletadas sobre unidades consumidoras de energia elétrica [1] [3] [4], principalmente aquelas processadas pelos sistemas de cadastro, solicitação de serviços, leitura de consumo e despacho de ordens de inspeção das empresas distribuidoras. Aprovado no ciclo de P&D Cemig ANEEL 2006/2007 e com resultados preliminares recém publicados, o projeto possibilita aos usuários da empresa escolher as melhores estratégias e clientes alvos, reduzindo assim os custos com inspeções mal sucedidas.

Metodologia

A metodologia do projeto consiste em duas fases. A primeira em realizar o agrupamento de dados multidimensionais de um repositório da empresa através do SOM e como resposta fornecer uma representação bidimensional topologicamente ordenada dos dados de maneira que aqueles com características intrínsecas similares tenderão a ficar em regiões vizinhas [6]. Uma vez que o algoritmo do SOM tenha convergido, o mapa de características calculado mostra características estatísticas importantes do espaço de entrada. Na segunda fase é construído um esquema de classificação supervisionado por meio de uma rede neural do tipo Perceptrons de Múltiplas Camadas para cada agrupamento formado com objetivo de identificar classes de consumidores irregulares e não-irregulares.

Resultados

Considerando apenas a variável “serviços solicitados pelo consumidor”, tais como débito automático, denúncia, emissão de segunda via de conta, entre outros, juntamente com informações de consumo e cadastro, a análise da matriz de distância unificada (U-matrix) e de agrupamentos através dos planos de componentes trouxe informações valiosas para apoio a tomada de decisão, principalmente aquelas relativas à detecção de perfis de unidades consumidoras que apresentam alguma irregularidade na medição. A Figura 1 apresenta a formação de 4 agrupamentos de consumidores cujas variáveis mais representativas para o problema foram a faixa de consumo (CLA), área geográfica (RAZ), rota de leitura (ROT) e atividade econômica (RAMO).



Percebe-se que a maior parte das instalações irregulares (P) concentra-se no agrupamento do canto inferior direito da U-matrix, em uma área geográfica (RAZ) e rota de leitura (ROT) bem definida. Os demais agrupamentos apresentam instalações regulares (N), porém dispersos em várias faixas de consumo e ramo de atividade. Duas grandes regiões dividem o mapa em instalações regulares, cujas variáveis CLA e RAMO a diferenciam. Inúmeras análises podem ser realizadas nesse sentido, fornecendo informações valiosas para apoio a tomada de decisão.

Figura 2 – Análise de desempenho da rede SOM

#	Mapa	Ano	Erro de quantização	Erro topográfico	RSME	Tempo CPU [som_make]	Tempo de Treinamento
1	[50,39]	2003-2007	0.261	0.096	0.3212	148.9030s	31s
2	[33,27]	2003	0.106	0.086	0.2998	19.7185s	6s
3	[35,26]	2004	0.101	0.061	0.6406	19.7497s	7s
4	[34,27]	2005	0.114	0.094	0.1965	21.5401s	7s
5	[32,25]	2006	0.092	0.048	0.2292	14.6329s	5s
6	[31,25]	2007	0.105	0.070	0.2134	13.5097s	4s

A Figura 2 apresenta as medidas de desempenho adotadas e informam que a melhor configuração da rede SOM para análise anual é treinamento batch, inicialização linear a partir do cálculo dos autovetores e autovalores dos dados de entrada, função de vizinhança tangente hiperbólica e mapa hexagonal com tamanho médio [33,26].

Referências

- [1] Eller, N.A. “Arquitetura de informação para o gerenciamento de perdas comerciais de energia elétrica”. Tese de Doutorado, Universidade Federal de Santa Catarina, Departamento de Engenharia de Produção e Sistemas, Florianópolis, 2003.
- [2] Fawcett, T.; Provost, F. Adaptive Fraud Detection. Data Mining and Knowledge Discovery 1, Netherlands: Kluwer Academic Publishers, 291-316, (1997).
- [3] Francisco, E. R., Petrielli, A. & Reina, C. S. “Segmentação comportamental de clientes para o setor elétrico”. São Paulo, In: Anais do III Congresso anual de tecnologias da informação da Fundação Getúlio Vargas, pp. 1-8, 2006.
- [4] Sperandio, M., Coelho, J., Queiroz, H., “Identificação de Agrupamentos de consumidores de Energia Elétrica através de SOM”. In: V SBQEE, Vol. 2, pp. 439-444, 2003.