

# SISTEMA INTELIGENTE PARA IDENTIFICAÇÃO DE FRAUDES EM REDES DE ENERGIA ELÉTRICA BASEADO EM LÓGICA FUZZY

EDUARDO WERLEY SILVA DOS ÂNGELOS\*, OSVALDO R. SAAVEDRA†, OMAR A. CARMONA CORTES‡

\**Universidade Federal do Maranhão  
Programa de Pós-Graduação em Engenharia de Eletricidade - PPGEE  
São Luís, Maranhão, Brasil*

†*Departamento de Engenharia de Eletricidade  
Universidade Federal do Maranhão (UFMA)  
Av. Portugal, S/N - Bacanga  
São Luís, MA, Brasil*

‡*Departamento Acadêmico de Informática  
Centro Federal de Educação Tecnológica do Maranhão (CEFET-MA)  
Av. Getúlio Vargas, 04 - Monte Castelo  
São Luís, MA, Brasil*

Emails: [eduardowerley@hotmail.com](mailto:eduardowerley@hotmail.com), [o.saavedra@ieee.org](mailto:o.saavedra@ieee.org), [omar@cefet-ma.br](mailto:omar@cefet-ma.br)

**Abstract**— In this paper we propose an intelligent approach based on Fuzzy Logic to help on mitigating non-technical losses in electric power networks, more precisely the ones linked to electricity fraud by costumers. By starting with the definition of some relevant source of electricity frauds, which are used as system's input variables, the relationships between them using linguistic if-then rules is stated. By analyzing customers' profiles taking into account the established variables, an individual index showing the chance of electricity fraud in such client is generated. The approach has been validated using a random data base, showing good performance in the fraudulent users process detection.

**Keywords**— Distribution Networks, Non-technical losses, Electricity Fraud Detection, Fuzzy Logic Systems.

**Resumo**— Neste artigo, propõe-se uma metodologia inteligente, baseada na Lógica Fuzzy, para auxiliar na redução das perdas comerciais em redes de energia elétrica relacionadas ao furto intencional de energia por parte dos consumidores. A partir do levantamento de possíveis causas motivadoras da prática de fraude — que serão as variáveis de entrada do sistema — são estabelecidas relações entre elas através de regras lingüísticas do tipo se-então. Analisando-se o perfil de consumo de determinado cliente sob a ótica das variáveis estabelecidas, é gerado então um índice representando a chance da ocorrência de fraudes. Para a validação do sistema foi criado e usado um banco de dados de consumidores, o que permitiu observar um bom desempenho no processo de detecção de fraudes.

**Palavras-chave**— Sistemas de Distribuição, Perdas Comerciais, Fraude de Energia, Sistemas Fuzzy.

## 1 Introdução

Associada à geração, transmissão e distribuição de energia elétrica sempre há uma parcela mínima de perdas. Estas perdas normalmente estão relacionadas à características naturais dos vários componentes do sistema, sendo que as mesmas ocorrem geralmente por dissipação de calor nos condutores. As perdas naturais são referenciadas pelo termo *perdas técnicas*. A partir do momento que as perdas de energia ultrapassam determinadas margens ótimas, elas passam a ser um fator primordial para a estabilidade operacional e financeira do sistema. As perdas financeiras, representadas pelo déficit entre a energia entregue e a energia comercializada são designadas pelo termo *perdas comerciais*, advindas principalmente de fraudes de energia, ligações clandestinas e inadimplência (Vidinich et al, 2007).

Enquanto a determinação da magnitude das perdas técnicas é feita diretamente, levando-se em consideração modelos matemáticos dos

diversos componentes da rede, a determinação das perdas comerciais não segue uma metodologia definida. Isto porque, neste último caso, ocorre a influência de inúmeros fatores humanos relacionados a questões sociais, culturais, econômicas e até políticas, onde a influência de cada uma delas varia conforme a região.

A principal forma de determinação das perdas comerciais se dá através de inspeções técnicas nas unidades consumidoras. Entretanto, a taxa de acerto na busca de fraudes ainda é baixa se comparada à magnitude total das perdas. Ainda que vários elementos tenham sido comprovadamente apontados como motivadores da prática de fraude, não existe ainda um sistema padronizado que tenha a capacidade de compilar o conhecimento adquirido e, ainda por cima, possa estabelecer diretrizes mais eficazes para combatê-la.

Neste trabalho, propõe-se a utilização da lógica fuzzy como estratégia para a modelagem dos processos de raciocínio que estão por trás

dos elementos motivadores do furto de energia, combinando diversas variáveis e suas correlações com a prática de fraude em um sistema único, simples e flexível.

A lógica fuzzy tem se mostrado eficiente na detecção de fraudes em ambientes comerciais como pode ser visto nos trabalhos de Pathak (Pathak et al., 2005), Phuong (Phuong et al., 2002), Kou (Kou et al., 2004) e Thang (Thang et al., 2006), pois pode trabalhar com dados imprecisos e até mesmo subjetivos quando coletados. Nesse contexto, vislumbrou-se a utilização da mesma técnica na detecção de fraudes em sistemas de energia.

O sistema desenvolvido visa auxiliar no processo de definição de prioridades para as inspeções técnicas, onde a partir do levantamento do perfil do cliente sob a ótica das variáveis de entrada do sistema, é inferido por um processo de raciocínio fuzzy a chance com que ele pode estar cometendo fraude.

Para alcançar seu objetivo, este artigo está dividido da seguinte forma: na Seção 2 apresenta-se uma visão geral das perdas de energia no Brasil e no mundo; na Seção 3 mostra-se a aplicação da lógica fuzzy para estruturar o sistema proposto; na Seção 4 realizam-se alguns testes sobre a viabilidade da metodologia utilizando um banco de dados fictício; e na Seção 5 são apresentadas as conclusões deste trabalho.

## 2 Perdas em Redes de Energia Elétrica

Segundo a Agência Internacional de Energia, as perdas de eletricidade nos sistemas de transmissão e distribuição mundiais durante o ano de 2004 alcançaram o patamar de 8,79% da energia total produzida naquele ano, ou seja, cerca de 1540 TW.h de acordo com a Tabela 1 (IEA, 2004).

| BLOCO                    | Geração (GW.h) | Perdas (GW.h) | Perdas (%) |
|--------------------------|----------------|---------------|------------|
| Pacífico (OECD)          | 1.729.598      | 81.785        | 4,73       |
| China e Hong Kong (OECD) | 2.236.733      | 143.614       | 6,42       |
| América do Norte (OECD)  | 4.997.072      | 339.937       | 6,80       |
| Europa(OECD)             | 3.468.931      | 242.898       | 7,00       |
| África                   | 539.427        | 59.758        | 11,08      |
| Oriente Médio            | 588.196        | 65.631        | 11,16      |
| Ex-União Soviética       | 1.379.568      | 176.490       | 12,79      |
| Europa (demais países)   | 197.040        | 25.775        | 13,08      |
| América Latina           | 875.473        | 146.478       | 16,73      |
| Ásia (exceto China)      | 1.518.952      | 258.361       | 17,01      |
| MUNDO - TOTAL            | 17.530.990     | 1.540.727     | 8,79       |

Tabela 1: Índices mundiais de perdas na distribuição em 2004

Percebe-se pela tabela mostrada que a distribuição dos índices de perdas por regiões mundiais não ocorre da mesma forma, sendo menor nos países desenvolvidos e substancialmente maior nos países em

desenvolvimento. Nestes últimos países, a ocorrência de elevados índices de perdas comerciais pode ser explicado devido a inadimplência, a falta de manutenção continuada e principalmente ao furto de energia (IEA, 2003).

No Brasil, a média histórica de perdas está estimada em 15%, tendo em 2004 alcançado o patamar de 16,85%, conforme ilustrado na Figura 1, onde é mostrado um comparativo de perdas entre os principais países em desenvolvimento (IEA, 2004).

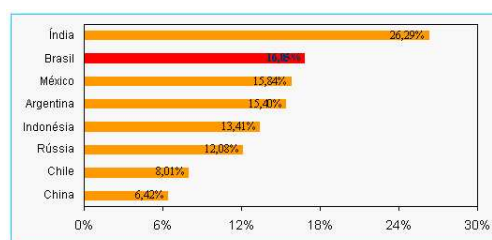


Figura 1: Perdas de energia em países em desenvolvimento

Utilizando o valor médio de aquisição da energia das distribuidoras (R\$ 80/MW.h), o índice de perdas no país em 2004 equivalerá em termos financeiros a cerca de 4 bilhões de reais. Já considerando o valor médio de venda de energia (R\$ 231,35), as perdas ocasionaram prejuízos da ordem de 11 bilhões de reais (Vidinich et al, 2007).

Os altos índices de perdas se refletem invariavelmente nas tarifas pagas pelos consumidores e no próprio sistema, no que se refere à geração e transmissão da energia adicional para compensar as perdas. Além desses fatores, há também o custo ambiental, tanto em termos de emissões de poluentes na atmosfera quanto em relação à infra-estrutura para acomodar o transporte da energia adicional.

### 2.1 Perdas Comerciais no Brasil

A Figura 2 representa a trajetória das perdas comerciais no Brasil, entre os anos 1997 e 2003 (ABRADEE, 2004), onde percebe-se uma tendência de alta das perdas nos últimos anos.

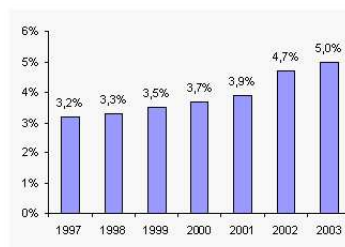


Figura 2: Trajetória das perdas comerciais no Brasil

O problema das perdas comerciais tem merecido nos últimos anos grande atenção por parte dos agentes do setor elétrico brasileiro. Para contornar essa questão, a Agência Nacional de Energia Elétrica (ANEEL) tem agido em dois campos: no campo regulatório, quando estabelece limites mínimos para os índices de perdas comerciais das concessionárias; e no campo administrativo, quando incentiva as concessionárias a buscarem estratégias e tecnologias que visem a redução do problema.

### 2.1.1 Causas

Diferentemente das perdas técnicas, as perdas comerciais são de difícil categorização, uma vez que são conseqüências de fenômenos altamente relacionados ao comportamento humano, variando ainda de região para região. Mesmo assim, nas áreas com elevados índices de perdas comerciais, existem certas características em comum, que podem ser associadas aos altos índices de perdas. São elas:

- *Furto de Energia:* Constitui-se no maior responsável pelos altos índices de perdas comerciais. Há duas estratégias principais de furto de energia: ligações diretas à rede (ligações clandestinas) e violações no medidor de energia. As ligações clandestinas se originam principalmente devido à ocupação desordenada em algumas regiões, onde não há fornecimento regular de energia, surgindo também por consumidores que se autoconectam à rede, mesmo diante da existência do fornecimento regular. Já as fraudes no medidor de energia ocorrem geralmente por manipulações nos seus terminais, componentes internos e instalações externas, fazendo com que seja registrado um consumo abaixo do real.
- *Violência e Ocupação Desordenada:* Dificultam o acesso das concessionárias de energia em algumas regiões (principalmente aquelas de grande *favelização*) o que impede as atividades de inspeção e corte de eletricidade, elevando as perdas comerciais (Portes, 2004).
- *Erros da Concessionária:* Associados a medição incorreta (falhas na leitura, falhas no medidor), ao faturamento irregular (unidades consumidoras sem cadastro ou taxadas incorretamente) e a terceirização dos serviços. (Vieiralves, 2005).
- *Outros fatores:* Questões culturais relacionadas ao desperdício de energia e a impunidade, e fatores econômicos associados a carga tributária do setor elétrico brasileiro e a inadimplência (ABRADEE, 2004).

### 2.1.2 Estratégias de combate às perdas

De uma maneira geral, a redução das perdas comerciais está centrada em gestões mais eficientes, com a realização de investimentos adequados e aplicação de estratégias que visem inibir a prática de fraudes.

A principal estratégia usada pelas distribuidoras no combate ao furto de energia tem sido a inspeção das unidades consumidoras. Quando detectada uma variação significativa no consumo de eletricidade, são acionadas equipes de inspeção para avaliar o problema. Entretanto, são muitos os casos de falhas na identificação de fraudes, acarretando um prejuízo da imagem das distribuidoras de energia perante a sociedade.

Ultimamente, algumas concessionárias têm investido também em novos sistemas de medição, como medidores eletrônicos, medição remota e medição nos principais alimentadores. Outras soluções técnicas têm sido tomadas, como a elevação do nível de redes secundárias, o uso de novas tecnologias para cabos condutores e detecção de fraudes nas instalações elétricas (Bazelatto et al, 2005).

Deve ser citada também a eficácia de medidas sociais no combate às perdas, como a conscientização sobre o uso correto da eletricidade e sobre as conseqüências da prática de fraude, o desenvolvimento de parcerias com vários setores da sociedade – aproximando a concessionária dos consumidores – e o estabelecimento de novos métodos de faturamento (AMPLA, 2005).

## 3 Metodologia para Detecção de Fraudes utilizando Lógica Fuzzy

A eficácia da identificação dos consumidores fraudadores de energia pode ser melhorada utilizando ferramentas inteligentes que possam correlacionar à prática de fraude seus elementos motivadores. No presente trabalho, utiliza-se a lógica fuzzy para modelar os processos de raciocínio que estão por trás das causas do furto de energia, através de uma metodologia flexível que pode integrar os conhecimentos adquiridos acerca da prática de fraudes e raciocinar sobre eles através de regras lingüísticas extraídas de especialistas.

### 3.1 Organização do sistema

O sistema desenvolvido correlaciona a influência de possíveis fatores de perdas comerciais em uma dada região a um índice final de perdas, baseando-se nos princípios do raciocínio aproximado modelados pela lógica fuzzy. Seu objetivo final é a melhoria da eficácia das inspeções técnicas através da prévia identificação dos consumidores com maior chance de fraude,

onde as informações a serem processadas por ele serão coletadas por leituristas, a partir da análise subjetiva e objetiva em cada unidade consumidora.

Em resumo, um sistema baseado em lógica fuzzy fundamenta-se nos *conjuntos fuzzy* e nas *variáveis lingüísticas*. Os conjuntos fuzzy, propostos por Lofti A. Zadeh (Zadeh, 1965), são uma generalização dos conjuntos clássicos, onde a pertinência de um dado elemento não é dual, podendo ser expressa por números reais no intervalo  $[0, 1]$ . Já as variáveis lingüísticas, também extensões da teoria fuzzy (Zadeh, 1975), fornecem meios para a representação de fenômenos mal-definidos ou imprecisos, sendo definidas como variáveis cujos valores são conjuntos fuzzy. Os conjuntos fuzzy quando associados às variáveis lingüísticas são representados pelas *funções de pertinência*.

A implementação do sistema proposto foi realizada com o auxílio do software FuzzyTech. O sistema está estruturado conforme o esquema mostrado na Figura 3, sendo composto por três etapas: *fuzzificação*, *inferência* e *defuzzificação*.

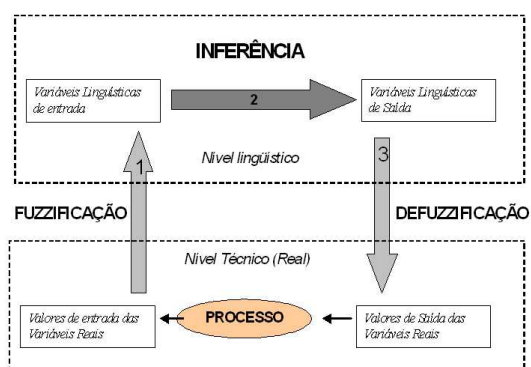


Figura 3: Organização de um sistema fuzzy

Durante a etapa de *fuzzificação* uma informação numérica, relacionada a uma determinada variável lingüística, é associada a um conjunto fuzzy composto por vários termos lingüísticos. Na etapa de *inferência*, o conjunto das informações de entrada são inter-relacionados a uma variável lingüística de saída através de regras do tipo *se-então*. Finalmente, no processo de *defuzzificação* o conjunto fuzzy na saída da etapa de inferência é associado a um valor final na variável lingüística de saída do sistema.

### 3.1.1 Variáveis Lingüísticas

Dentre os vários fatores motivadores das perdas comerciais, foram escolhidos os de maior correlação com a prática de fraudes e relacionados em cinco variáveis lingüísticas: *consumo*, *ocupação*, *perfil\_região*, *peso\_conta* e *satisfação*. Estas são as variáveis de entrada do sistema, sendo as duas primeiras avaliadas

diretamente pelas leituras e as demais a partir de respostas dadas pelos clientes. Para a saída do sistema foi formulada a variável *indice\_fraude*, que avalia a chance de um dado consumidor estar cometendo uma fraude de energia.

Para cada variável lingüística também foram elaboradas funções de pertinência, associando valores numéricos de entrada a graus de pertinência no intervalo unitário. Igualmente, foram definidas a extensão e os métodos de avaliação dos valores numéricos de cada variável como apresentado a seguir:

**Consumo:** Esta variável tem a função de representar a relação entre o consumo atual do cliente e sua respectiva média de consumo nos últimos seis meses. Sua magnitude será dada em valores por unidade (p.u.), de acordo com a Equação 1:

$$W(pu) = \frac{W_{atual}}{\bar{W}} \quad (1)$$

onde,  $W_{atual}$  representa o consumo atual do cliente (em KW.h) e  $\bar{W}$  é a média do consumo dos últimos 6 meses.

A função de pertinência modelada para a variável *consumo* é ilustrada na Figura 4.

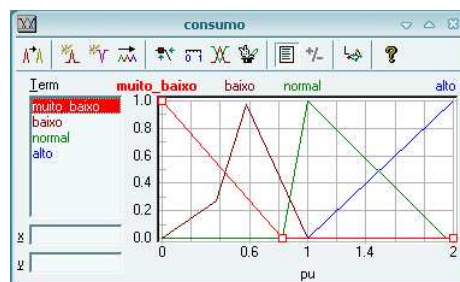


Figura 4: Modelo da variável *consumo*

**Ocupação:** Avalia a densidade de ocupação do imóvel; uma estimativa da quantidade de pessoas por área construída. A ocupação será dada em valores de 1 a 5, da seguinte forma: *ocupação muito baixa*=1; *baixa*=2; *normal*=3; *alta*=4 e *muito alta*=5. O modelo desta variável está ilustrado na Figura 5.

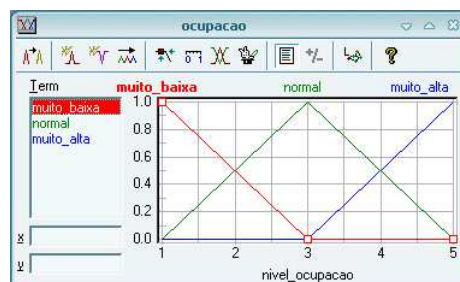


Figura 5: Modelo da variável *ocupação*

**Perfil\_região:** Caracteriza o perfil da região do imóvel, em termos de proximidade de favelas, urbanização e qualidade das instalações elétricas. Será dada em valores de 1 a 5, da seguinte forma: perfil *baixo*=1; *médio-baixo*=2; *médio*=3; *médio-alto*=4 e *alto*=5. O modelo desta variável está ilustrado na Figura 6.

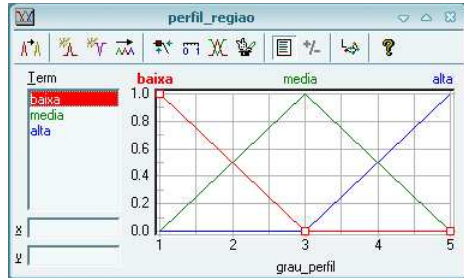


Figura 6: Modelo da variável *perfil\_região*

**Peso\_conta:** Tem a função de avaliar o peso da conta de energia no orçamento mensal do cliente e será informada aos leituristas pelos consumidores. As informações serão em graus de 1 a 3, da seguinte forma: peso da conta *baixo*=1; *normal*=2; *alto*=3. O modelo das funções de pertinência é dado na Figura 7.

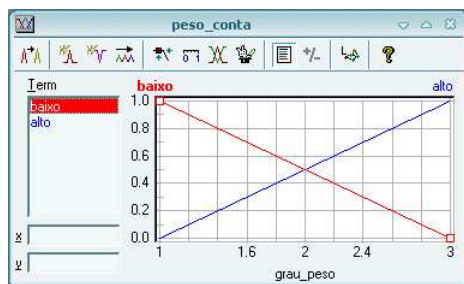


Figura 7: Modelo da variável *peso\_conta*

**Satisfação:** Representa o grau de satisfação do cliente quanto aos serviços prestados pela concessionária, sendo avaliado em notas de 0 a 10, conforme o modelo da Figura 8.

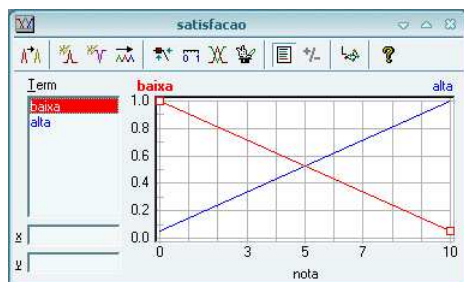


Figura 8: Modelo da variável *satisfação*

**Índice\_fraude:** Variável de saída do sistema, que indica um índice de 0 a 1 representando a

chance de um dado consumidor estar fraudando energia, conforme o modelo apresentado na Figura 9.

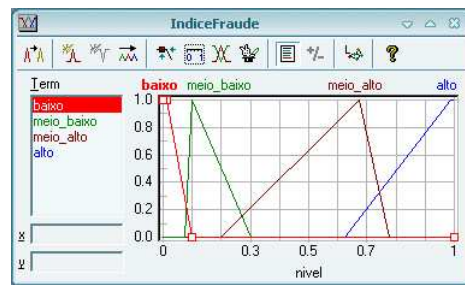


Figura 9: Modelo da variável *índice\_fraude*

### 3.1.2 Base de Conhecimento

O conhecimento registrado no sistema proposto está na forma de regras lingüísticas do tipo *se x = A então y = B*, onde *x* e *y* são os dados de entrada dos termos lingüísticos (conjuntos fuzzy) *A* e *B*, respectivamente. Inicialmente foram estabelecidas 4 hipóteses gerais de interação entre as variáveis estabelecidas, com base no conhecimento registrado na bibliografia.

**Hipótese 1** *A variação do consumo provoca uma variação inversamente proporcional no índice de fraude, exceto quando acompanhada pela variação da ocupação.*

**Hipótese 2** *A taxa de crescimento do índice de fraude com a diminuição do consumo é maior em regiões com perfil mais baixo.*

**Hipótese 3** *Um peso da conta elevado, somado a um consumo reduzido, com a mesma ocupação, eleva o grau de fraude.*

**Hipótese 4** *Um índice de satisfação baixo, aliado a um consumo baixo, com a mesma ocupação, eleva o grau de fraude.*

Após a definição das hipóteses gerais, formulou-se a base de regras do sistema, conforme mostrado na Tabela 2. A tabela de regras não possui todas as possíveis combinações entre as variáveis dada a não aplicabilidade de alguns casos. A base de regras foi estabelecida de maneira hierárquica, isto é, as variáveis subjetivas *peso\_conta* e *satisfação* só terão relevância quando o conhecimento registrado pelas demais variáveis for incerto.

## 4 Avaliação do Sistema

O ideal, para validar o sistema proposto, seria dispor de dados reais de consumo de energia nas concessionárias. No entanto, dada a dificuldade

Tabela 2: Base de Regras do Sistema

| #  | consumo     | ocupação    | perfil_regiao | peso_conta | satisfação | DoS   | IndiceFraude |
|----|-------------|-------------|---------------|------------|------------|-------|--------------|
| 1  | muito_baixo | muito_baixa | baixa         | -          | -          | 1,000 | meio_baixo   |
| 2  | muito_baixo | muito_baixa | media         | -          | -          | 1,000 | meio_baixo   |
| 3  | muito_baixo | muito_baixa | alta          | -          | -          | 1,000 | meio_baixo   |
| 4  | muito_baixo | normal      | baixa         | alto       | baixa      | 1,000 | alto         |
| 5  | muito_baixo | normal      | media         | alto       | baixa      | 1,000 | alto         |
| 6  | muito_baixo | normal      | alta          | alto       | baixa      | 1,000 | alto         |
| 7  | muito_baixo | normal      | baixa         | baixo      | alta       | 0,898 | alto         |
| 8  | muito_baixo | normal      | media         | baixo      | alta       | 0,898 | alto         |
| 9  | muito_baixo | normal      | alta          | baixo      | alta       | 0,898 | alto         |
| 10 | muito_baixo | muito_alta  | baixa         | -          | -          | 1,000 | alto         |
| 11 | muito_baixo | muito_alta  | media         | -          | -          | 1,000 | alto         |
| 12 | muito_baixo | muito_alta  | alta          | -          | -          | 1,000 | alto         |
| 13 | baixo       | muito_baixa | baixa         | -          | -          | 1,000 | meio_baixo   |
| 14 | baixo       | muito_baixa | media         | -          | -          | 1,000 | meio_baixo   |
| 15 | baixo       | muito_baixa | alta          | -          | -          | 1,000 | meio_baixo   |
| 16 | baixo       | normal      | alta          | baixo      | alta       | 1,000 | meio_alto    |
| 17 | baixo       | normal      | media         | baixo      | alta       | 0,898 | meio_alto    |
| 18 | baixo       | normal      | alta          | alto       | baixa      | 0,797 | meio_alto    |
| 19 | baixo       | normal      | media         | alto       | baixa      | 0,500 | meio_alto    |
| 20 | baixo       | normal      | baixa         | baixo      | alta       | 0,398 | meio_alto    |
| 21 | baixo       | normal      | baixa         | alto       | baixa      | 0,297 | meio_alto    |
| 22 | baixo       | muito_alta  | baixa         | -          | -          | 1,000 | alto         |
| 23 | baixo       | muito_alta  | media         | -          | -          | 1,000 | alto         |
| 24 | baixo       | muito_alta  | alta          | -          | -          | 1,000 | alto         |
| 25 | normal      | muito_baixa | baixa         | -          | -          | 1,000 | baixo        |
| 26 | normal      | muito_baixa | media         | -          | -          | 1,000 | baixo        |
| 27 | normal      | muito_baixa | alta          | -          | -          | 1,000 | baixo        |
| 28 | normal      | normal      | baixa         | -          | -          | 1,000 | meio_baixo   |
| 29 | normal      | normal      | media         | -          | -          | 1,000 | meio_baixo   |
| 30 | normal      | normal      | alta          | -          | -          | 1,000 | meio_baixo   |
| 31 | normal      | muito_alta  | alta          | baixo      | baixa      | 1,000 | meio_alto    |
| 32 | normal      | muito_alta  | alta          | baixo      | alta       | 1,000 | meio_alto    |
| 33 | normal      | muito_alta  | alta          | alto       | baixa      | 1,000 | meio_alto    |
| 34 | normal      | muito_alta  | alta          | alto       | alta       | 1,000 | meio_alto    |
| 35 | normal      | muito_alta  | media         | baixo      | baixa      | 0,703 | meio_alto    |
| 36 | normal      | muito_alta  | media         | baixo      | alta       | 0,703 | meio_alto    |
| 37 | normal      | muito_alta  | media         | alto       | baixa      | 0,703 | meio_alto    |
| 38 | normal      | muito_alta  | media         | alto       | alta       | 0,703 | meio_alto    |
| 39 | normal      | muito_alta  | baixa         | baixo      | baixa      | 0,297 | meio_alto    |
| 40 | normal      | muito_alta  | baixa         | baixo      | alta       | 0,297 | meio_alto    |
| 41 | normal      | muito_alta  | baixa         | alto       | baixa      | 0,297 | meio_alto    |
| 42 | normal      | muito_alta  | baixa         | alto       | alta       | 0,297 | meio_alto    |
| 43 | alto        | normal      | baixa         | -          | -          | 1,000 | baixo        |
| 44 | alto        | normal      | media         | -          | -          | 1,000 | baixo        |
| 45 | alto        | normal      | alta          | -          | -          | 1,000 | baixo        |

de aquisição de tais dados, elaborou-se um banco de dados fictício composto por 30 perfis de consumo, como registrado na Tabela 3. Formularam-se também para cada caso os valores numéricos das variáveis lingüísticas de entrada.

Tomando-se como entrada os valores contidos no banco de dados, simulou-se o sistema com os métodos de defuzzificação *média dos máximos (MOM)*, *centro dos máximos (COM)* e *centro de área (COA)*. Para o processo de inferência, utilizou-se os o operador (mínimo) para a implicação e conjunção e o operador *máximo* para a disjunção. Os resultados estão registrados na Tabela 3 nas colunas **Índice de Fraude**.

#### 4.1 Análises

As simulações utilizando o método de defuzzificação *média dos máximos (MOM)* apresentaram pouca diversidade nos valores numéricos, o que pode dificultar o processo de classificação de clientes fraudulentos. Os dados mais consistentes foram obtidos com os métodos *centro de área (CoA)* e *centro dos máximos (CoM)*, sendo que esta última estratégia apresentou maior faixa de variação e associatividade com as regras e funções de pertinência elaboradas.

##### 4.1.1 Teste das Hipóteses

A comprovação das hipóteses definidas para o sistema será realizada a partir da comparação entre perfis de consumo da Tabela 3, tendo como base os índices de fraude gerados pelo método CoM.

- **Hipótese 1:** Levando-se em consideração os perfis 1, 2 e 3, observa-se que, fixando-se um dado perfil de consumo (*perfil 1*: 1 p.u.; índice de fraude 0,1000), quando o consumo por unidade se eleva (*perfil 2*: 1,5 p.u.) o índice se reduz para 0,0519. Para o *perfil 3*, com um consumo de 0,25 p.u., o índice de fraude simulado foi de 0,9082. Estas análises comprovam a hipótese 1.

Analisando-se somente a variável ocupação, percebe-se que o *perfil 6* (consumo em p.u. 0,25; ocupação *baixa*) tem índice de fraude (0,5650) menor que o *perfil 7* (0,9147), que possui perfil idêntico ao dele exceto com a ocupação *alta*. Uma ocupação alta, neste caso, sugere uma contradição entre o consumo a quantidade de pessoas no imóvel, aumentando o índice de fraude final.

- **Hipótese 2:** Analisando-se os consumidores 9, 10 e 11, observa-se que para o cliente com perfil da região informado *alto (perfil 11)*, o índice (0,8071) é menor do que o cliente de perfil informado *médio* (0,8155) e do que o

cliente de perfil *baixo* (0,8790), o que está de acordo com a hipótese estabelecida.

Para consumos normais, quanto mais baixo é o perfil da região, maior será o índice de fraude final, conforme conclui-se a partir dos perfis 2 e 3 da tabela 2.

- **Hipótese 3:** Tomando-se os perfis 14 e 15, observa-se que para um consumo *baixo* e ocupação *normal*, o peso na conta informado *baixo (perfil 14)* fornece um índice de fraude (0,8155) menor para o caso do peso na conta alto (*perfil 15*) – índice 0,87. Isto comprova a definição da hipótese 3, que estabelece que o peso da conta reduzido diminui o índice de fraude final.

Um peso da conta alto no orçamento familiar pode indicar uma dificuldade em se manter financeiramente o mesmo nível mensal de consumo. Se em determinado mês o consumo diminui, permanecendo inalterada a ocupação do imóvel, tendo sido previamente informado um peso na conta *alto*, isto pode indicar a existência de fraude.

- **Hipótese 4:** Observando-se os perfis 3 e 8, nota-se que para uma satisfação maior (*perfil 3*: nota 10), o índice de fraude (0,9082) é menor do que no caso de uma satisfação menor (*perfil 8*: nota 2,5) – índice de fraude 0,9148 – o que comprova a hipótese 4.

Ainda que mínima, a diferença dos índices de fraude neste caso pode ser particularmente útil em um processo de classificação de clientes, na definição de prioridades para a realização de inspeções técnicas.

##### 4.1.2 Considerações Adicionais

Para clientes em total dissonância com a prática de fraude, o índice de fraude simulado é próximo de 0,10. Índices abaixo deste valor podem indicar um consumo de energia excessivamente alto.

Observa-se que, para consumos abaixo ou acima da média na taxa de 10%, em condições normais, os índices de fraude não são altos, representando a margem deixada nas funções de pertinência para oscilações normais no consumo mensal.

Nota-se também que o sistema tem a capacidade de detectar dissonâncias tanto para altas quanto para baixas médias de consumo. Isto permitiria, a priori, sua aplicação em grandes consumidores de energia. Entretanto o sistema é mais eficaz para análise de perfis residenciais de consumo, tendo suas variáveis especificamente formuladas para este cenário.

Tabela 3: Banco de Dados Fictício e Índices de Fraude Simulados

| Perfil | W <sub>atual</sub> | W   | W(pu) | Ocupação    | Perf_Região | Peso_Conta | Satisf. | Índice Fraude |        |        |
|--------|--------------------|-----|-------|-------------|-------------|------------|---------|---------------|--------|--------|
|        |                    |     |       |             |             |            |         | (MOM)         | (COM)  | (COA)  |
| 1      | 400                | 400 | 1,00  | normal(3)   | média(3)    | normal(2)  | 5       | 0,1000        | 0,1000 | 0,1583 |
| 2      | 600                | 400 | 1,50  | normal(3)   | média(3)    | normal(2)  | 5       | 0,0062        | 0,0519 | 0,1153 |
| 3      | 100                | 400 | 0,25  | normal(3)   | média(3)    | normal(2)  | 10      | 0,9937        | 0,9082 | 0,7604 |
| 4      | 800                | 800 | 1,00  | normal(3)   | alta(5)     | normal(2)  | 5       | 0,1000        | 0,1000 | 0,1583 |
| 5      | 150                | 150 | 1,00  | normal(3)   | md_baixa(2) | normal(2)  | 5       | 0,1000        | 0,1000 | 0,1583 |
| 6      | 100                | 400 | 0,25  | baixa(2)    | média(3)    | normal(2)  | 5       | 0,1000        | 0,5650 | 0,5971 |
| 7      | 100                | 400 | 0,25  | alta(4)     | média(3)    | normal(2)  | 10      | 0,9937        | 0,9147 | 0,7683 |
| 8      | 100                | 400 | 0,25  | normal(3)   | média(3)    | normal(2)  | 2,5     | 0,9937        | 0,9148 | 0,7682 |
| 9      | 200                | 400 | 0,50  | normal(3)   | baixa(1)    | normal(2)  | 10      | 0,9937        | 0,8790 | 0,7270 |
| 10     | 200                | 400 | 0,50  | normal(3)   | média(3)    | normal(2)  | 10      | 0,6750        | 0,8155 | 0,6627 |
| 11     | 200                | 400 | 0,50  | normal(3)   | alta(5)     | normal(2)  | 10      | 0,6750        | 0,8071 | 0,6549 |
| 12     | 230                | 230 | 1,00  | alta(4)     | baixa(1)    | normal(2)  | 5       | 0,1000        | 0,2315 | 0,3273 |
| 13     | 450                | 450 | 1,00  | alta(4)     | alta(5)     | normal(2)  | 5       | 0,1000        | 0,3875 | 0,4398 |
| 14     | 150                | 300 | 0,50  | normal(3)   | md_baixa(2) | baixo(1)   | 10      | 0,6750        | 0,8155 | 0,6627 |
| 15     | 150                | 300 | 0,50  | normal(3)   | md_baixa(2) | alto(3)    | 10      | 0,9937        | 0,8877 | 0,7365 |
| 16     | 220                | 220 | 1,00  | mt_baixa(1) | média(3)    | normal(2)  | 8       | 0,0062        | 0,0062 | 0,0339 |
| 17     | 330                | 200 | 1,65  | mt_alta(5)  | baixa(1)    | alto(3)    | 1,5     | 0,0062        | 0,0905 | 0,2529 |
| 18     | 60                 | 800 | 0,08  | mt_baixa(1) | alta(5)     | alto(3)    | 6       | 0,1000        | 0,1000 | 0,1583 |
| 19     | 1000               | 500 | 2,00  | normal(3)   | md_alta(4)  | normal(2)  | 5       | 0,0062        | 0,0062 | 0,0339 |
| 20     | 300                | 300 | 1,00  | normal(3)   | alta(5)     | alto(3)    | 9,5     | 0,1000        | 0,1000 | 0,1583 |
| 21     | 50                 | 50  | 1,00  | normal(3)   | baixa(1)    | alto(3)    | 8       | 0,1000        | 0,1000 | 0,1583 |
| 22     | 50                 | 665 | 0,08  | normal(3)   | md_alta(4)  | normal(2)  | 5       | 0,9937        | 0,9603 | 0,8267 |
| 23     | 600                | 500 | 1,20  | normal(3)   | alta(5)     | baixo(1)   | 7       | 0,1000        | 0,0811 | 0,1443 |
| 24     | 400                | 500 | 0,80  | normal(3)   | alta(5)     | baixo(1)   | 3,5     | 0,6750        | 0,6961 | 0,5649 |
| 25     | 325                | 300 | 1,08  | alta(4)     | média(3)    | normal(2)  | 10      | 0,1000        | 0,3089 | 0,3995 |
| 26     | 150                | 165 | 0,91  | baixo(2)    | alta(5)     | normal(2)  | 5       | 0,0062        | 0,1621 | 0,2988 |
| 27     | 50                 | 665 | 0,08  | baixo(2)    | baixa(1)    | normal(2)  | 5       | 0,1000        | 0,5498 | 0,6091 |
| 28     | 355                | 357 | 0,99  | normal(3)   | alta(5)     | alto(3)    | 0       | 0,1000        | 0,1109 | 0,1768 |
| 29     | 200                | 244 | 0,82  | mt_baixa(1) | média(3)    | baixo(1)   | 0       | 0,1000        | 0,1000 | 0,1583 |
| 30     | 600                | 520 | 1,15  | baixo(2)    | md_baixa(2) | alto(3)    | 10      | 0,0062        | 0,0531 | 0,1168 |

## 5 Conclusões

Este artigo apresentou um sistema inteligente para detecção de fraudes baseado em lógica fuzzy. Os objetivos definidos para a metodologia proposta foram alcançados com a classificação dos prováveis fraudadores de energia, levando-se em consideração as variáveis estabelecidas e suas correlações à prática de fraude. Com a definição de prioridades nas inspeções em busca de fraudes de energia, as concessionárias poderão otimizar recursos humanos e materiais, bem como terem minimizados os desgastes perante a comunidade nos casos de erros de avaliação.

Deve ser salientado que a verificação propriamente dos fraudadores de energia se dá somente através de inspeções técnicas, com a emissão de termos de irregularidades que sejam certificados e validados por órgãos credenciados. Mesmo assim, esta última etapa não encerra o processo, na medida em que deve existir ainda um entendimento entre consumidores fraudulentos e concessionárias quanto a forma de pagamento da energia desviada, bem como existir mecanismos legais eficientes para fazer com que os acordos sejam cumpridos e, caso não sejam cumpridos, haver mecanismos operacionais para manter os fraudadores com o fornecimento de energia interrompido até a regularização da situação.

No sistema proposto, a forma de raciocínio sobre a ocorrência ou não de fraude de energia para um determinado perfil de cliente está inteiramente condicionada à base de regras formulada para o sistema. Assim, se em determinada localidade as hipóteses estabelecidas

não forem verdadeiras, ou forem irrelevantes, a metodologia ainda poderá ser aplicada, necessitando-se a princípio somente que seja alterada sua base de regras, de modo a se compatibilizar com o novo raciocínio, o que garante uma enorme portabilidade e flexibilidade para o sistema.

É importante considerar que a eficácia da metodologia dependerá inteiramente da confiabilidade das respostas ou valores numéricos das variáveis lingüísticas. Para que haja uma melhor eficiência final, é necessário que sejam estabelecidos critérios bem definidos acerca da aquisição dos dados para análise, bem como investir no treinamento dos funcionários e na tecnologia a ser usada no processo de registro dos dados.

Futuros aperfeiçoamentos no sistema poderiam fazer com que o mesmo levasse em consideração, na definição das variáveis lingüísticas, outros fatores, como: curvas de carga diárias ou sazonais; média do consumo reativo; levantamento do histórico de fraudes da região analisada; variação no consumo de água; e existência de débitos em sistemas de proteção ao crédito; passagem pela polícia e comparação de perfis semelhantes (panificadoras, hotéis, motéis, frigoríficos, etc). Não há limites para a quantidade de variáveis do sistema.

Semelhantemente, poderiam ser utilizados métodos inteligentes, como redes neurais ou algoritmos genéticos, para se otimizar, além da forma e quantidade das funções de pertinência, a quantidade dos termos lingüísticos utilizados, o tipo do método de defuzzificação e operadores



lógicos. Uma outra sugestão para trabalhos futuros seria implementar uma estratégia inteligente de realimentação, de modo que os erros ou acertos obtidos pudessem tornar o sistema mais eficiente.

### Referências

- ABRADEE (2004). Reflexos da inadimplência no setor de energia elétrica, *In: V Seminário Jurídico. Goiânia-GO*. Disponível em <http://www.abradee.org.br/Downloads/Juridicos/>. Acesso em 12 dez. 2006.
- AMPLA (2005). Relatório de sustentabilidade. Disponível em <http://www.ampla.com/files/>. Acesso em 12 dez. 2006.
- Bazelatto et al, G. (2005). Um dispositivo eletrônico para detectar fraude de energia elétrica, *In: VI Seminário Internacional de Metrologia Elétrica. Rio de Janeiro – RJ*.
- IEA (2003). *World Energy investment Outlook 2003*, International Energy Agency & Europe Copper Institute.
- IEA (2004). IEA electricity statistics. Disponível em <http://www.iea.org/Textbase/stats/prodresult.asp?PRODUCT=Electricity/Heat>. Acesso em 15 maio 2007.
- Kou, Y., Lu, C.-T., Sirirat, S. and Huang, Y.-P. (2004). Survey of fraud detection techniques, *IEEE International Conference on Networking, Sensing and Control*, pp. 21–23.
- Pathak, J., Vidyarthi, N. and Summers, S. L. (2005). A fuzzy-based algorithm for auditors to detect elements of fraud in settled insurance claims, *Managerial Auditing Journal* **20**(6): 632–644.
- Phuong, N. H., Santiprabhob, P., Thang, C. and Masayuki, A. (2002). A fuzzy consultation system for computer configurations, *International Conference InTech/VJFuzzy*, pp. 137–142.
- Portes, J. H. F. (2004). Impacto da violência nas atividades da light – RJ, *In: I Workshop Furto/Fraude de Energia e Roubo de Condutores e Equipamentos*. Disponível em [http://www.abradee.org.br/Downloads/1\\_Worshop\\_Furtos/light.pdf](http://www.abradee.org.br/Downloads/1_Worshop_Furtos/light.pdf). Acesso em 12 dez. 2006.
- Thang, C., Toan, P. Q., Cooper, E. W. and Kamei, K. (2006). Application of soft computing to tax fraud detection in small businesses, *IEEE International Conference on Communications and Electronics*, pp. 402–407.
- Vidinich et al, R. (2007). Considerações sobre as perdas na distribuição de energia elétrica no brasil, *Revista Brasil Engenharia*. num. 580.
- Vieiralves, E. X. (2005). *Proposta de uma metodologia para avaliação das perdas comerciais dos sistemas elétricos: O caso manaus*, Dissertação de mestrado, Faculdade de Engenharia Mecânica: Universidade Estadual de Campinas.
- Zadeh, L. A. (1965). Fuzzy sets, *Information and Control* **8**: 338–353.
- Zadeh, L. A. (1975). The concept of a linguistic variable and its application to approximate reasoning, *Information Sciences* **3**.