**Instituto de Pesquisas Tecnológicas do Estado de São Paulo**

**Charles Pereira Niza**

# Detecção de fraudes de consumo de energia em *Smart grids* utilizando a teoria de Dempster-Shafer

**São Paulo**

**2014**

Charles Pereira Niza

# Detecção de fraudes de consumo de energia em *Smart grids* utilizando a teoria de Dempster-Shafer

Dissertação de Mestrado apresentada ao Instituto de Pesquisas Tecnológicas do Estado de São Paulo – IPT, como parte dos requisitos para a obtenção do título de Mestre em Engenharia de Computação.

Data da aprovação _____/_____/_____

_____
Prof. Dr. Eduardo Takeo Ueda (Orientador)
Centro Universitário Senac

Membros da Banca Examinadora:

Prof. Dr. Eduardo Takeo Ueda (Orientador)
Centro Universitário Senac

Prof. Dr. Alexandre José Barbieri de Sousa (Membro)
IPT – Instituto de Pesquisas Tecnológicas do Estado de São Paulo

Prof. Dr. Marcos Antonio Simplicio Junior (Membro)
Poli-USP

Charles Pereira Niza

# Detecção de fraudes de consumo de energia em *Smart grids* utilizando a teoria de Dempster-Shafer

Dissertação de Mestrado apresentada ao Instituto de Pesquisas Tecnológicas do Estado de São Paulo – IPT, como parte dos requisitos para a obtenção do título de Mestre em Engenharia de Computação.

Área de Concentração: Redes de Computadores

Orientador: Prof. Dr. Eduardo Takeo Ueda

São Paulo
Novembro/2014

Ficha Catalográfica
Elaborada pelo Departamento de Acervo e Informação Tecnológica – DAIT
do Instituto de Pesquisas Tecnológicas do Estado de São Paulo - IPT

N737d **Niza, Charles Pereira**
Detecção de fraudes de consumo de energia em Smart grids utilizando a teoria de Dempster-Shafer. / Charles Pereira Niza. São Paulo, 2014.
93p.

Dissertação (Mestrado em Engenharia de Computação) - Instituto de Pesquisas Tecnológicas do Estado de São Paulo. Área de concentração: Redes de Computadores.

Orientador: Prof. Dr. Eduardo Takeo Ueda

1. Detecção de fraude 2. Consumo de energia elétrica 3. Teoria de Dempster-Shafer 4. Tese I. Ueda, Eduardo Takeo, orient. II. IPT. Coordenadoria de Ensino Tecnológico III. Título

14-09 CDU 004.7(043)

Aos meus pais, José Carlos e Djanira, aos meus filhos, Charles e Kevin, à minha esposa Juliana e ao meu irmão Patrick que sempre me incentivaram e apoiaram.

## AGRADECIMENTOS

Agradeço ao meu orientador, Prof. Dr. Eduardo Takeo Ueda pelo apoio, incentivo e, sobretudo por sua confiança na minha capacidade de desenvolver este trabalho.

Agradeço aos professores Alexandre Barbieri e Marcos Simplicio pelas valiosas contribuições dadas ao longo do desenvolvimento do trabalho.

Agradeço a professora Edit pelas importantes orientações que foram dadas ao longo da disciplina de Metodologia da Pesquisa.

Agradeço aos funcionários da Secretaria Acadêmica, em especial, o Sr. Adilson Feliciano pela presteza.

Agradeço ao senhores João, Ronaldo, Airton e Juliano da Gerência de Recuperação de energia da CPFL Paulista pela valiosa ajuda.

Agradeço ao Prof. Bussab, Diretor dos Cursos de Informática da UNINOVE pelo apoio total e irrestrito.

Agradeço ao meu colega, Silvio Rocha pela amizade.

## RESUMO

As redes elétricas inteligentes foram concebidas com o objetivo de aumentar a qualidade e a confiabilidade de sistemas de transmissão e distribuição de energia e possibilitar a integração de novas fontes de energia limpa ao sistema elétrico. A adoção dessas redes permitirá às concessionárias de energia e aos consumidores mudarem a forma como produzem e consomem energia, propiciando o estabelecimento de novas modalidades tarifárias e novos comportamentos de consumo. Com a implantação das redes elétricas inteligentes, os medidores de consumo eletromecânicos serão substituídos por medidores inteligentes, que podem ser monitorados, lidos e comandados, remotamente, pela concessionária. Nesse cenário, o problema objeto de estudo desta dissertação é como garantir que as informações coletadas e enviadas para a distribuidora pelos medidores inteligentes de uma determinada unidade consumidora sejam, de fato, confiáveis e que o instrumento que as referenciam não tenha sido adulterado. Assim, em caso de suspeita de fraude, a distribuidora poderia executar os procedimentos de verificação e de inspeção em campo. Nesse sentido, o trabalho tem por objetivo aplicar um mecanismo de detecção de fraudes de consumo de energia elétrica utilizando a Teoria Matemática de Dempster-Shafer. A efetividade do mecanismo de detecção de fraudes aplicado foi realizada por meio de simulações utilizando uma amostragem estratificada obtida a partir de uma base de dados fornecida por uma concessionária de energia elétrica contendo dados reais de medições de consumo de energia elétrica de unidades consumidoras já inspecionadas e identificadas a priori como sendo regulares ou irregulares.



Palavras-chave: smart grid, medidores inteligentes, detecção de fraudes, Dempster-Shafer

## ABSTRACT

### Fraud detection of power consumption for Smart Grids using Dempster-Shafer theory

Smart grids are designed with the goal of increasing the quality and reliability of transmission and distribution systems and enable the integration of new sources of clean energy to the electrical system. The adoption of these networks will allow utilities and consumers to change the way we produce and consume energy, leading to the establishment of new tariff arrangements and new consumer behaviors. With the deployment of smart grids, electromechanical consumer meters will be replaced by smart meters, which can be tracked, scanned and controlled remotely by the utility. In this scenario, the problem subject of this dissertation is to ensure that the information collected and sent to the distributor of the smart meters a particular consumer unit are in fact reliable and that the reference instrument that has not been tampered with. Thus, in the case of suspected fraud, the distributor could implement procedures for verification and field inspection. In this sense, the thesis aims to implement a mechanism for detecting fraud in electricity consumption using the Dempster-Shafer Theory. The effectiveness of the proposed fraud detection mechanism was performed by simulations using a stratified sample obtained from a database provided by an electric utility containing actual measurement data of electricity consumption of consumer units already inspected and identified a priori as being regular or irregular.



Keywords: smart grid, smart meters, fraud detection, Dempster-Shafer

## Lista de Ilustrações



.

## Lista de Tabelas

## Lista de Abreviaturas e Siglas

| | |
|---|---|
| ANEEL | Agência Nacional de Energia Elétrica |
| AUC | *Area Under Curve* |
| $CO_2$ | Dióxido de Carbono |
| DNP3 | *Distributed Network Protocol* |
| GPRS | *General Packet Radio Service* |
| GSM | *Groupe Spécial Mobile* (sistema global para comunicações móveis) |
| HAN | *Home Area Network* |
| IC | Intervalo de Confiança |
| IEDs | *Intelligent Electronic Devices* (dispositivos eletrônicos inteligentes) |
| IEEE | *Institute of Electrical and Electronics Engineers* |
| IPCA | Índice de Perda Comercial do Alimentador |
| ISA | Índice de Suspeita da Área |
| NAN | *Neighborhood Area Network* |
| NIST | *National Institute of Standards and Technology* |
| PAN | *Personal Area Network* |
| PC | Perfil de Comportamento |
| ROC | *Receiver Operating Characteristics* |
| RTUs | *Remote Terminal Units* (Unidades Terminais Remotas) |
| TIC | Tecnologia da Informação e Comunicação |
| WAN | *Wide Area Network* |

**SUMÁRIO**

# 1 INTRODUÇÃO

A adoção das redes elétricas inteligentes permitirá às concessionárias de energia e aos consumidores mudarem a forma como produzem e consomem energia. Além de aumentar a qualidade e a confiabilidade do sistema de transmissão e distribuição de energia, integrar novas fontes de energia limpa ao sistema e promover o uso racional e eficiente de energia elétrica, a adoção das redes elétricas inteligentes propiciará o estabelecimento de novas modalidades tarifárias e novos comportamentos de consumo.

Em termos práticos, as redes elétricas inteligentes propiciam automação, monitoração e integração de todos os instrumentos e equipamentos dos sistemas de geração, distribuição, medição e armazenamento de energia. Assim, os medidores e instrumentos de campo realizam suas funções de maneira autônoma, permitindo que a rede elétrica, por meio de análises e diagnósticos feitos em tempo real, se reconfigure automaticamente para atender, de forma otimizada, às necessidades do sistema elétrico.

Com a implantação das redes elétricas inteligentes, os atuais medidores de consumo de energia serão substituídos, gradativamente, por medidores inteligentes, que poderão ser monitorados, lidos e comandados remotamente pela concessionária.

A utilização das redes elétricas inteligentes permitirá, entre outras vantagens, ajustar o consumo da energia, de modo que seja possível ao usuário se beneficiar de períodos com tarifas reduzidas e evitar períodos nos quais os preços são mais altos. Além disso, também será possível realizar o corte de fornecimento de energia de uma residência ou grupo de residências específicos, visando à melhoria do desempenho da rede, em situações críticas.

No entanto, devido à complexidade, à diversidade de instrumentos e equipamentos, à descentralização do sistema e ao elevado volume de informações transmitidas é necessário estabelecer mecanismos capazes de garantir a confidencialidade, a integridade e a disponibilidade das informações

registradas e transmitidas entre os dispositivos do consumidor e os medidores e os concentradores da rede. Como resultado, propiciar uma proteção mínima contra ataques do tipo *man-in-the-middle*, forjamento de informação e de negação de serviço. Para Khurana et al. (2010), os dados coletados pelos medidores inteligentes devem ser precisos e confiáveis.

Este trabalho apresenta um mecanismo de detecção de fraudes de consumo de energia elétrica utilizando dois métodos (duas abordagens) operando em paralelo: uma baseada em análise diferencial, centrada no consumo de energia elétrica (kWh) da unidade consumidora e que detecta mudanças significativas nos padrões de consumo no decorrer do tempo e outra baseada em análise global, centrada na observação do comportamento global das unidades consumidoras, utilizando como parâmetro, o degrau de consumo. As evidências de fraude determinadas por meio destas duas abordagens serão combinadas utilizando-se a Teoria de Evidência de Dempster-Shafer.

## 1.1 Motivação

De acordo com Toledo (2012), é necessário conferir mecanismos de irretratabilidade que assegurem que a origem da informação seja realmente o instrumento que é dado como referência e vice-versa, quando a informação ou o comando é proveniente da distribuidora.

Tão importante quanto autenticar as partes envolvidas é garantir a confiança das informações que serão coletadas pelos medidores inteligentes, instalados junto às unidades consumidoras e repassadas para a distribuidora. Evita-se, assim, além do comprometimento da confiabilidade e da eficiência operacional do sistema elétrico como um todo, perdas financeiras decorrentes dos desvios e furtos de energia (FARIA, 2012).

Neste cenário, portanto, o problema objeto de estudo desta dissertação é como garantir que as informações coletadas pelos medidores inteligentes de uma

determinada unidade consumidora e enviadas para a distribuidora sejam, de fato, confiáveis, bem como a integridade do instrumento que as referenciam. Assim, em caso de suspeita de fraude, a distribuidora poderia disparar procedimentos de verificação e de inspeção em campo.

Visando encontrar um valor capaz de representar e mensurar a confiança dentro de determinados contextos, existem alguns trabalhos que apontam possíveis direções a ser exploradas, dentre elas a da utilização de modelos de confiança computacional.

Albuquerque (2008) propõe a criação e validação de um modelo de confiança voltado para sistemas distribuídos. O trabalho de Benzi (2011) apresenta um modelo de confiança para um sistema de gerenciamento de conteúdo de comércio eletrônico, com base em requisitos de confiança elencados segundo a aplicabilidade a esse ambiente. O estudo realizado por Duarte et al. (2008) apresenta e analisa um modelo de confiança para redes móveis *ad hoc*. Fraga et al. (2013) propõem um modelo de confiança para a composição de serviços *Web*.

Bograd (2012) propõe um modelo de confiança e reputação baseado na utilização de sistemas multiagentes no contexto da geração distribuída, no qual, o consumidor poderá fornecer energia elétrica gerada a partir de fontes renováveis para a rede ou outros consumidores.

O estudo realizado por Pradhan et al. (2011) propõe a utilização de uma abordagem baseada em confiança e reputação para detectar variações atípicas do consumo de energia reportadas pelos medidores inteligentes. O estudo foi realizado com base na análise e comparação dos dados de consumo das unidades consumidoras em relação aos padrões de consumo individual e de outras unidades consumidoras dentro da mesma categoria de consumo.

O trabalho realizado por Matei et al. (2012) apresenta um modelo de confiança baseado em sistemas multiagentes para verificação da precisão dos dados de consumo registrados pelos medidores inteligentes e enviados para

concessionária, a fim de propiciar uma proteção adicional contra forjamento de dados de consumo em virtude de adulteração de equipamentos ou ataques cibernéticos.

Diferentemente dos trabalhos estudados, seja em função da utilização de padrões de fraude já conhecidas ou em virtude da utilização da análise da variação do padrão de consumo da própria unidade consumidora ou mesmo em relação a outras unidades consumidoras da mesma categoria na identificação de fraude, o presente trabalho prevê a utilização de dois métodos de detecção de suspeita de fraude operando em paralelo: um baseado em análise diferencial, e outro baseado em análise global, cujas evidências obtidas serão combinadas utilizando-se a Teoria Matemática de Dempster-Shafer.

Apesar de ser usual a representação da confiança em sistemas computacionais, não existe um consenso sobre sua aplicação e utilização de modo generalizado, uma vez que a aplicação da confiança se dá dentro de cenários e contextos específicos, cujas particularidades e restrições inviabilizam a utilização dos modelos propostos em outras perspectivas.

## 1.2 Objetivo

O presente trabalho tem por objetivo propor e aplicar um mecanismo de detecção de fraudes de consumo de energia elétrica utilizando a Teoria Matemática de Dempster-Shafer baseando-se no método de Kovach (2011).

A efetividade do mecanismo de detecção de fraudes proposto será avaliada por meio da validação do modelo proposto e da análise dos resultados obtidos a partir da realização de simulações utilizando uma amostragem estratificada gerada a partir de uma base de dados fornecida por uma concessionária de energia elétrica contendo dados reais de medições de consumo de energia elétrica de unidades consumidoras já inspecionadas e identificadas a priori como sendo regulares ou irregulares.

## 1.3 Contribuições

O presente trabalho tem como objetivo, além de contribuir com o equacionamento do problema de detecção de fraude de consumo de energia elétrica, servir como uma alternativa para a combinação de evidências de fraudes determinadas por outros modelos de detecção de fraudes já utilizados pelas concessionárias, propiciando maior diversificação de modelos a fim de ampliar a visão dos especialistas sobre o problema da detecção de fraude de consumo de energia elétrica.

A aplicação do mecanismo de detecção de fraudes de consumo de energia proposto neste trabalho contribuirá com o aumento da precisão das informações coletadas pelos medidores inteligentes instalados junto às unidades consumidores e enviadas para a rede distribuidora, assegurando a integridade da origem da informação. Essa contribuição permitirá que as concessionárias reduzam a probabilidade de fraudes e de perdas financeiras decorrentes delas, bem como melhorem a eficiência do sistema elétrico como um todo.

O aumento do grau de confiança das informações coletadas e transmitidas em uma rede elétrica inteligente, decorrente da efetividade do mecanismo proposto neste trabalho contribuirá, oportunamente, com a viabilização da introdução de novos modelos tarifários, que permitirão ao consumidor beneficiar-se de tarifas diferenciadas, de acordo com seu perfil de consumo.

Além de equacionar a questão central da detecção de fraudes dos dados de consumo coletados pelos medidores inteligentes, o presente trabalho propiciará uma contribuição ao tema por meio do estabelecimento de um conjunto mínimo de requisitos capaz de representar e mapear confiança em redes elétricas inteligentes. Possibilita-se ainda que futuros trabalhos contemplem outros aspectos relacionados à confiabilidade da informação, como a identificação dos instrumentos de medição inteligentes e a proteção das trocas de informação entre os instrumentos de medição e os concentradores da rede.

Assim, como resultado deste trabalho, pretende-se avançar no estado da arte referente ao aspecto da detecção de fraudes em S*mart grids* – essencial para a confiabilidade da operação dos sistemas de redes elétricas inteligentes.

## 1.4 Método de trabalho

Pesquisa descritiva e exploratória visando ao desenvolvimento e aplicação de um mecanismo capaz de detectar fraudes de consumo de energia elétrica. Tendo em vista o objetivo que se pretende alcançar neste trabalho, serão desenvolvidas as seguintes atividades:

a) revisão da literatura e análise dos principais métodos de detecção de fraudes, buscando identificar contribuições e aspectos que possam ser objeto de aprofundamento. Definição dos principais conceitos e fundamentos relacionados ao tema deste trabalho, como geração e armazenamento distribuído de energia, tomadas, eletrodomésticos e medidores inteligentes de energia elétrica.

b) desenvolvimento e aplicação de um mecanismo de detecção de fraudes de consumo de energia elétrica utilizando a Teoria Matemática de Dempster-Shafer que seja capaz de gerar um escore de suspeita de fraude com base nas evidências determinadas por dois módulos operando em paralelo.

c) realização de simulações e validação do mecanismo proposto a fim de verificar a eficiência do mecanismo proposto para detectar fraudes no registro de medições de consumo de energia elétrica. Serão realizadas simulações utilizando uma amostragem estratificada gerada a partir de uma base de dados fornecida por uma concessionária de energia elétrica contendo dados reais de medições de consumo de energia elétrica de unidades consumidoras de um município do interior de São Paulo, já inspecionadas e identificadas a priori como sendo regulares ou irregulares.

d) análise crítica dos resultados decorrentes da aplicação do mecanismo de detecção de fraudes do consumo de energia elétrica proposto neste trabalho. A

representação dos resultados obtidos nas simulações será feita por meio de tabelas e gráficos e sua correspondente análise.

## 1.5 Organização do Trabalho

Essa dissertação está organizada em 7 capítulos, como segue:

O Capítulo 1, “INTRODUÇÃO”, discorre sobre a motivação, o objetivo, as contribuições e o método empregado no trabalho.

O Capítulo 2, *“SMART GRIDS* E DETECÇÃO DE FRAUDE*”*, apresenta uma revisão dos principais conceitos, fundamentos e trabalhos relacionados à *Smart grids* e os dois principais métodos de detecção de fraudes que servirão de base para o entendimento do trabalho.

O Capítulo 3, sobre “TEORIA DE EVIDÊNCIA DE DEMPSTER-SHAFER”, apresenta uma revisão da literatura sobre a Teoria Matemática de Evidências de Dempster-Shafer, buscando identificar contribuições e aspectos relevantes dentro da perspectiva da detecção de fraudes.

O Capítulo 4, “DETECÇÃO DE FRAUDES DE CONSUMO DE ENERGIA UTILIZANDO A TEORIA DE DEMPSTER-SHAFER” apresenta os detalhes da proposta de mecanismo de detecção de fraude de consumo de energia elétrica utilizando a teoria de evidência de Dempster-Shafer.

O Capítulo 5, “VALIDAÇÃO DA PROPOSTA”, apresenta as principais considerações acerca da validação da proposta.

O Capítulo 6, “ANÁLISE DOS RESULTADOS”, apresenta uma análise crítica dos resultados obtidos nas simulações realizadas decorrentes da aplicação do mecanismo de detecção de fraudes proposto.

O Capítulo 7, “CONCLUSÕES”, apresenta as principais conclusões acerca do trabalho e sugestões para trabalhos futuros.

# 2 *SMART GRIDS* E DETECÇÃO DE FRAUDES

Este capítulo tem por objetivo apresentar uma revisão dos principais conceitos, fundamentos e trabalhos relacionados à *Smart grids* e os dois principais métodos de detecção de fraudes que servirão de base para o entendimento do trabalho.

## 2.1 Definição de *Smart grids*

As redes elétricas inteligentes surgiram em resposta às necessidades do setor elétrico de aumentar a qualidade e a confiabilidade do sistema de transmissão e distribuição de energia e de integrar fontes de energia renováveis à matriz energética. A adoção das redes elétricas inteligentes permite às concessionárias de energia a promoção do uso racional e a redução das perdas de energia elétrica, atualmente existentes, contribuindo, dessa forma, com o desenvolvimento sustentável e a preservação do meio ambiente.

Em termos práticos, a adoção de S*mart grids* propiciará a automação, monitoração e integração de todos os instrumentos e equipamentos dos sistemas de geração, distribuição, medição e armazenamento de energia. Nesse contexto, os medidores de consumo de energia elétrica e demais instrumentos de campo passam a ser capazes de realizar suas funções de maneira autônoma, permitindo assim, que a rede elétrica, por meio de análises e diagnósticos em tempo real, reconfigure-se automaticamente para atender de forma otimizada as necessidades do sistema elétrico.

S*mart grid* pode ser definido como:

> "... a transição de um modelo caracterizado pela geração centralizada e redes passivas de distribuição de energia elétrica para um modelo caracterizado pelo gerenciamento e controle da geração realizada pelas diferentes fontes que passarão a fazer parte das chamadas microredes conectadas de forma distribuída à rede elétrica...", segundo Oliveira e Júnior (2012).

Essa transição é alcançada por meio da adoção de tecnologias da informação e comunicação e da convergência entre a infraestrutura de geração, transmissão e distribuição de energia elétrica e a infraestrutura de comunicação de dados, interligando os chamados IEDs (*Intelligent Electronic Devices*) e permitindo a troca de informações e ações de controle entre os diversos segmentos da rede elétrica.

Segundo Oliveira e Júnior (2012), a transição do atual modelo de sistema de geração, transmissão e distribuição de energia elétrica para o modelo baseado em S*mart grids* ocorrerá de forma gradativa, sendo que o primeiro passo será a substituição dos atuais medidores de consumo de energia elétrica por medidores inteligentes.

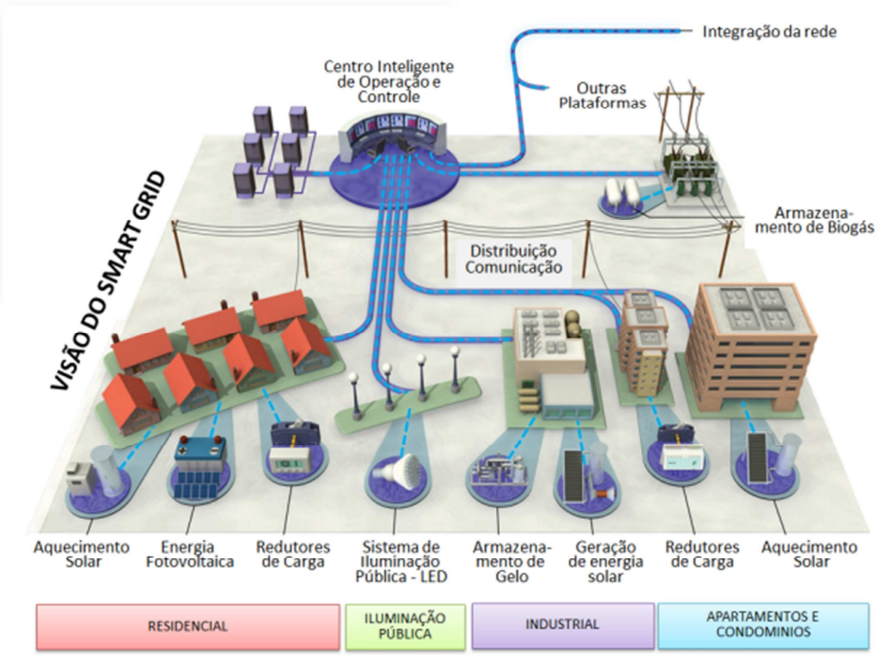
Os medidores inteligentes, instalados junto às unidades consumidoras, poderão ser monitorados, lidos e comandados remotamente pela concessionária. Com a adoção das redes elétricas inteligentes será possível monitorar o consumo de energia elétrica de cada cliente em tempo real, de tal forma que, eventuais falhas sejam percebidas imediatamente. Além disso, a introdução de tecnologias da informação para gerenciamento da rede poderá favorecer a redução das perdas e o aumento da eficiência dos sistemas como um todo (SMART E-NERGY, 2010).

A utilização desses medidores inteligentes fornecerá condições para a introdução de novos modelos tarifários, capazes de mudar o comportamento de toda a cadeia consumidora de energia elétrica (OLIVEIRA; JÚNIOR, 2012).

Com a adoção de S*mart grids* será possível agregar à rede elétrica, novas fontes de geração de energia, dentre as quais, fontes de energia limpa, como a energia eólica e a solar. Com as novas fontes de geração de energia conectadas diretamente à rede ou às unidades consumidoras será possível aumentar a capacidade de integração de geração distribuída das centrais de microgeração nas redes elétricas, principalmente, em redes de baixa tensão – conceito de *microgrid* (OLIVEIRA; JÚNIOR, 2012).

Um dos principais benefícios da adoção de S*mart grids* é justamente essa possibilidade de oferecer e integrar diversas fontes de energia, permitindo que os consumidores escolham quais fontes querem utilizar e em quais horários, a chamada geração distribuída. Pode-se converter o excedente de energia produzido em descontos na conta de luz, além de devolvê-la para a rede, o que diminui os riscos de sobrecarga na rede. A Figura 1 apresenta uma visão geral de um S*mart grid*.

Figura 1 - Visão geral de um S*mart grid*




Fonte: nMentors (2013)

## 2.2 Benefícios da Adoção de *Smart grids*

A adoção de *smart grids* prevê benefícios tanto para as concessionárias de energia quanto para os consumidores. A fim de se apresentar uma visão geral sobre alguns dos principais benefícios da adoção de S*mart grids*, de acordo com os trabalhos estudados, dividiu-se o tratamento do assunto sob dois aspectos: o

primeiro aborda os benefícios sob a perspectiva dos consumidores e o segundo, a partir da visão das concessionárias.

### 2.2.1 Benefícios da adoção de S*mart grids* sob a perspectiva dos consumidores

Sob a perspectiva dos consumidores, a adoção de S*mart grids*, permitirá que eles acompanhem, em tempo real, as informações sobre o consumo de energia elétrica e o nível de qualidade da energia recebida, possibilitando assim, melhor planejamento e controle dos gastos com energia elétrica, por meio da adequação do consumo ao orçamento doméstico, uma vez que será possível o deslocamento de carga, principalmente dos horários de pico, para períodos cujas tarifas poderão ser reduzidas (OLIVEIRA; JÚNIOR, 2012).

### 2.2.2 Benefícios da adoção de S*mart grids* sob a perspectiva das concessionárias

Sob o ponto de vista das concessionárias, a adoção de S*mart grids* possibilitará a detecção e a correção automática de falhas na rede, em tempo real, o que agilizará a realização de manutenção preventiva e monitoramento da qualidade do fornecimento de energia, o suporte à geração e ao armazenamento distribuído de energia, o fornecimento detalhado de informações sobre o serviço prestado, bem como, a redução de perdas técnicas e melhoria na eficiência operacional.

Além dos benefícios citados, a adoção de S*mart grids* tem por objetivo prover alta confiabilidade e qualidade da energia elétrica fornecida, otimização da utilização dos ativos do sistema elétrico, minimização dos custos de operação e manutenção do sistema, capacidade de autodiagnostico e auto recuperação da rede elétrica frente a qualquer problema detectado no circuito elétrico, resistência

a ataques cibernéticos, capacidade de interconexão de grande variedade de fontes de geração na forma distribuída e opções de armazenamento de energia. (BROWN, 2008)

Conclui-se, portanto, que a adoção de S*mart grids***,** tanto na perspectiva das concessionárias, quanto dos consumidores, representará uma mudança de paradigma na forma de lidar com os sistemas de geração, transmissão e distribuição de energia elétrica.

## 2.3 Panorama da Adoção de *Smart grids* no Brasil

No Brasil, o primeiro teste de viabilidade das redes elétricas inteligentes vem sendo realizado pela concessionária de energia elétrica portuguesa EDP em Aparecida, município paulista de 35.000 habitantes, desde 2011 (EXAME, 2014).

Em Minas Gerais, a Cemig pretende substituir os medidores eletromecânicos por medidores inteligentes de 8000 unidades consumidoras da região de Sete Lagoas, município vizinho a Belo Horizonte, até abril de 2014 (EXAME, 2014).

Já a AES Eletropaulo, a maior concessionária do país em faturamento e consumo de energia, com 6,5 milhões de consumidores, pretende substituir os atuais 60000 medidores de energia elétrica, analógicos, de todos os clientes de Barueri, por medidores inteligentes até 2015 (EXAME, 2014).

No Brasil, a partir de março de 2014, os consumidores de baixa tensão, como residências e pequenos estabelecimentos comerciais poderão ser beneficiados pela chamada tarifa branca, ou seja, a cobrança de tarifas diferenciadas de acordo com o horário em que a energia elétrica for consumida. Esse é um dos aspectos que tem impulsionado a disseminação das redes elétricas inteligentes na Europa.

## 2.4 Visão geral do sistema de geração, transmissão e distribuição de energia elétrica

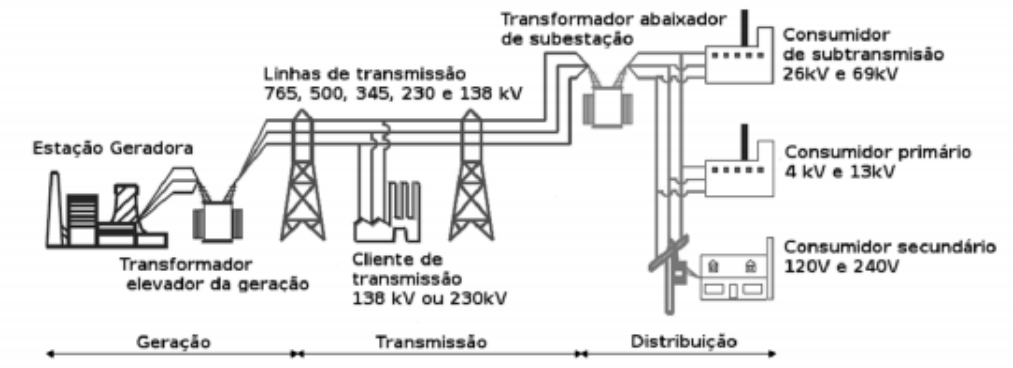
A energia elétrica é transportada das usinas de geração, sejam elas, hidrelétricas, termelétricas, eólicas ou termonucleares até os centros urbanos através das linhas das redes de transmissão de energia elétrica. Essas redes, além das linhas de transmissão, propriamente ditas, são compostas por um conjunto de cabos, torres e subestações de transformação equipadas com transformadores e equipamentos de controle e proteção (ABRADEE, 2014).

Como as plantas de geração de energia elétrica, geralmente, se encontram localizadas distantes das unidades consumidoras, a energia elétrica é transmitida das geradoras até os centros de distribuição locais, por meio de linhas de alta tensão. A energia elétrica é transmitida das geradoras até os centros de distribuição em alta tensão a fim de se reduzir as perdas elétricas em função da dissipação de energia nas linhas de transmissão (GUIMARÃES et al., 2013).

> “Em linhas gerais, a energia elétrica é conduzida através do sistema de transmissão, desde a sua origem na geração até subestações transformadoras de distribuição, que são responsáveis por distribuí-la através do sistema de redes de distribuição” (TOLEDO, 2012).

O equipamento responsável, tanto pela elevação quanto pelo rebaixamento do nível de tensão elétrica, em uma subestação é o transformador. O esquema básico de funcionamento do sistema elétrico é mostrado na Figura 2.

Figura 2 - Sistema elétrico básico




Fonte: Guimarães et al.(2013, p. 110)

As subestações de transmissão são instalações, geralmente, localizadas próximas às centrais de geração de energia elétrica. Seu objetivo é elevar o nível da tensão elétrica para que a energia elétrica seja transportada até os centros de distribuição locais.

Além de transformadores, uma subestação de transmissão dispõe de equipamentos de manobra, como chaves seccionadoras e disjuntores, equipamentos de medição e de proteção como relés, fusíveis e para-raios.

Já as subestações de distribuição, usualmente, localizadas nos grandes centros, próximo às zonas de consumo, são responsáveis pelo rebaixamento do nível de tensão elétrica para ser distribuída em média tensão.

Assim como ocorre em uma subestação de transmissão, os principais componentes de uma subestação de distribuição são transformadores, equipamentos de manobra, como chaves seccionadoras e disjuntores, equipamentos de medição e de proteção contra curtos-circuitos e descargas atmosféricas (ABRADEE, 2014).

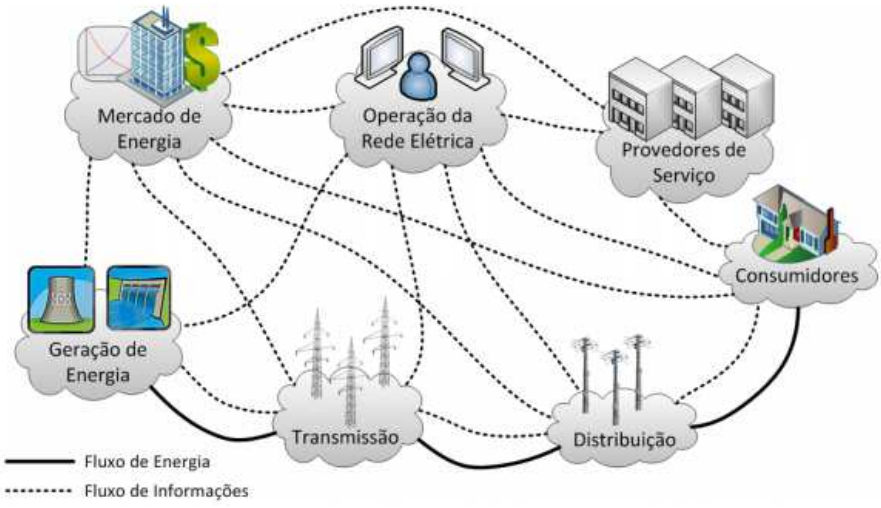
## 2.5 Características e componentes das redes elétricas inteligentes

As redes elétricas inteligentes têm por objetivo aumentar a confiabilidade, a eficiência e a qualidade do sistema elétrico como um todo. Segundo Guimarães et al. (2013):

> "As redes elétricas inteligentes propõem maior automação das subestações utilizando Dispositivos Eletrônicos Inteligentes (*Intelligent Devices – IEDs*) e Unidades Terminais Remotas (*Remote Terminal Units – RTUs*) para aumentar a capacidade de controle e monitoramento de dados".

O NIST (*National Institute of Standards and Technology*) propõe um modelo conceitual para redes elétricas inteligentes, composto por sete domínios de atores: geração de energia, transmissão, distribuição, consumidores, operação da rede elétrica, provedores de serviços e mercado de energia (NISTIR 7628, 2010). A Figura 3 apresenta uma visão geral desses domínios e a comunicação entre eles.

Figura 3 - Domínios das redes elétricas inteligentes




Fonte: Guimarães et al.(2013, p. 111)

De acordo com o modelo conceitual para redes elétricas inteligentes, proposto pelo NIST, o domínio de geração de energia é constituído pelas plantas de geração e armazenamento (NISTIR 7628, 2010).

Segundo Guimarães et al. (2013), “Os domínios de transmissão e distribuição são constituídos basicamente das subestações e das linhas de transmissão de energia...”. Sendo que o domínio de distribuição também é responsável por coletar os dados de consumo obtidos pelos medidores inteligentes, instalados junto às unidades consumidoras.

O domínio de consumidores contempla os aspectos relacionados ao consumo, armazenamento e à geração de energia em pequena escala. Os domínios de mercado e de provedores de serviços são, respectivamente, responsáveis pelo balanceamento da oferta e da demanda de energia elétrica e pela operação de serviços terceirizados.

Finalmente, o domínio de operação tem por objetivo propiciar a comunicação e a troca de informações entre os demais domínios da rede a fim de garantir o controle e a operação confiáveis da rede (GUIMARÃES et al., 2013). Os fluxos de informação entre os domínios são bidirecionais e, segundo Guimarães et al. (2013), “...englobam todos os domínios para garantir a interoperabilidade dos diversos serviços das redes elétricas inteligentes.”

### 2.5.1 Infraestrutura de comunicação

Quanto à abrangência, as redes de comunicação podem cobrir pequenas, médias e longas distâncias e podem interconectar milhões de equipamentos com ou sem restrição de capacidade de processamento (TOLEDO, 2012).

> Uma *HAN* (*Home Area Network)* “... corresponde à rede de comunicação formada entre o concentrador/interface com o consumidor e seus dispositivos...”. (TOLEDO, 2012). “Podem ainda, existir redes de menor abrangência, chamadas de *PAN* (*Personal Area Network)*, em geral, associadas a um indivíduo e aos seus próprios dispositivos...” (TOLEDO, 2012).

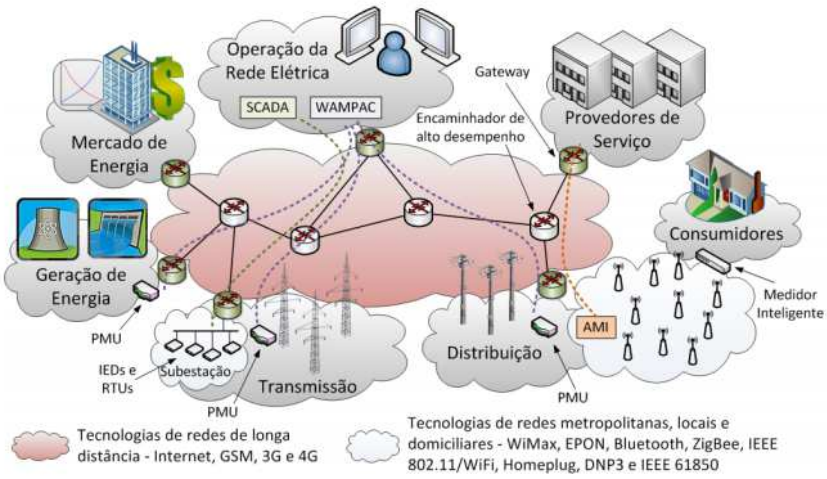
Esses dispositivos, geralmente, são conectados entre si por meio de protocolos de comunicação de curta distância, como o *Bluetooth*. O nível de abrangência, imediatamente superior, correspondente à rede formada pelos

medidores e os pontos de concentração de dados e é chamado de *NAN* (*Neighborhood Area Network)*. Uma *NAN* pode abranger inúmeras residências, quarteirões e até mesmo bairros.

A rede que conecta os concentradores da distribuidora é chamada de *WAN* (*Wide Area Network*). Uma *WAN* pode abranger grandes áreas. De acordo com TOLEDO (2012) “Todas essas redes podem operar de forma isolada e com certa autonomia.” As definições do fluxo de dados que atravessam os perímetros de cada rede, bem como, os requisitos de segurança que devem ser adotados em cada nível são determinantes para a garantia da confiabilidade das redes elétricas inteligentes (TOLEDO, 2012).

O sistema de comunicação das redes elétricas inteligentes prevê a utilização de padrões que possam aproveitar a disponibilidade preexistente na indústria de eletrodomésticos e de automação residencial, a fim de potencializar a sua adoção. Sistemas como: 802.11, *Bluetooth*, *Zigbee*, *Homeplug*, *Mesh*, *GSM/GPRS*, 3G e 4G podem ser escolhidos conforme a abrangência da rede e especificidades de cada região, bem como os padrões específicos utilizados pelas subestações de energia elétrica, como *DNP3* e IEEE 61850. A Figura 4 mostra as redes de comunicação entre os domínios (GUIMARÃES et al., 2013).

Figura 4 - Redes de comunicação entre os domínios




Fonte: Guimarães et al.(2013, p. 112)

### 2.5.2 Medidores inteligentes

O medidor inteligente é uma evolução dos medidores eletromecânicos. De acordo com Smart Grid Light (2014) “Além de permitir o envio automático de dados de medição para a concessionária, permitirá ao consumidor receber informações avançadas sobre seu consumo e fornecimento de energia.” Os medidores inteligentes além de fornecerem as funcionalidades metrológicas básicas, também são capazes de estabelecer canais de comunicação e disponibilizar informações sobre consumo para os usuários por meio da Internet.

Os medidores inteligentes são equipados com *displays* que exibem aos consumidores, informações sobre o consumo e emissões de $CO_2$ equivalentes à energia consumida diária, semanal e mensalmente, além das informações acumuladas e simuladas.

Com a adoção da tarifa branca no país, prevista para 2014, os consumidores poderão acompanhar pelo *display,* em tempo real, as informações sobre o consumo de energia elétrica e o valor das tarifas em um determinado instante, o que permitirá que gerenciem o próprio consumo de energia elétrica, de acordo com o valor da tarifa cobrada.

> “A esses medidores poderão ser associados mostradores remotos inteligentes que, além das funcionalidades supracitadas, permitirão ao consumidor acompanhar seu gasto com energia elétrica de forma amigável, através de textos, gráficos, imagens e cores”. SMART GRID LIGHT (2014)

O cliente poderá ainda, segundo Oliveira e Júnior (2012), “...estabelecer metas de consumo e receber alertas, toda vez que as estimativas de consumo projetarem que a meta poderá ser ultrapassada.”

Ao adequar seu comportamento de consumo de energia aos horários de tarifas menores, o cliente poderá se beneficiar reduzindo seus gastos com energia. A Figura 5 apresenta o modelo de medidor inteligente que será utilizado pela Companhia Energética de Minas Gerais (Cemig) em substituição dos 8.000

medidores eletromecânicos no município de Sete Lagos, em Minas Gerais, até abril de 2014 (PÁGINA SUSTENTÁVEL, 2013).

Figura 5 – Modelo de Medidor Inteligente


Fonte: Página Sustentável (2013)

## 2.5.3 Tomadas inteligentes

A micromedição, por meio das tomadas inteligentes, permitirá o chaveamento de cargas específicas que, associadas à adoção de tarifas diferenciadas, propiciará, de forma automática o deslocamento de carga ao longo do dia e a diminuição do desperdício de energia elétrica (deslocamento de consumo).

As tomadas inteligentes ajudarão a identificar o consumo real das cargas a elas ligadas a fim de promover o uso racional da energia elétrica junto à unidade consumidora por meio do chaveamento de cargas. Dessa forma, as tomadas inteligentes às quais equipamentos como TV, DVDs, decodificadores, micro-ondas e carregadores de bateria de celulares estiverem ligados, poderão ser programadas para desligarem completamente esses equipamentos durante o período, no qual, não estiverem sendo utilizados (da meia-noite as seis hora da manhã, por exemplo), eliminando assim, o consumo de energia proveniente do modo *stand-by.* Essa medida poderia representar uma economia de até 10% do consumo total de energia de uma residência.

De forma análoga, tomadas associadas a equipamentos de alto consumo poderão ser configuradas pelo cliente para funcionarem, prioritariamente, nos períodos em que a tarifa cobrada for mais baixa, propiciando dessa forma, benefícios tanto para o cliente quanto para o sistema elétrico como um todo (TOLEDO, 2012).

### 2.5.4 Armazenamento distribuído de energia

A adoção de S*mart grids* prevê também a produção de energia distribuída e integrada, transformando os clientes da concessionária em produtores. Para isso, as redes elétricas inteligentes preveem a integração de fontes renováveis de energia e armazenamento distribuído, incluindo veículos elétricos e híbridos recarregáveis. Sistemas de geração e armazenamento distribuídos instalados em uma unidade consumidora permitirão ao cliente utilizar o excedente em horários nos quais a tarifa cobrada é mais alta ou converter esse excedente em descontos na conta de luz. Segundo Toledo (2012), “Tal energia pode ser utilizada tanto para compensações de consumo em horário de ponta, aliviando o sistema de distribuição, quanto para alimentar cargas prioritárias em casos de interrupções no fornecimento de energia.”.

## 2.6 O problema da detecção de fraudes em S*mart grids*

A detecção de fraudes tem por objetivo identificar uma fraude antes que ela seja consumada. “... A detecção de fraudes entra em ação quando a prevenção não consegue evitar a fraude...” (KOVACH, 2011).

A AES Eletropaulo, maior concessionária do país em termos de faturamento e consumo de energia, estima que cerca de 4% da energia distribuída se perca em função de práticas ilícitas, como adulteração de medidores e ligações clandestinas. Essa quantidade seria suficiente para abastecer, por exemplo, o

município de Santo André, na Grande São Paulo, com 670.000 habitantes, por dez meses (EXAME, 2014).

No sentido de validar a fidelidade e a inviolabilidade dos dados de consumo coletados pelos medidores, de tal forma, que a concessionária, em caso de suspeita de tentativa de fraude, possa disparar procedimentos de inspeção em campo, o presente trabalho propõe o desenvolvimento e a aplicação de um mecanismo de detecção de fraude no consumo de energia utilizando duas abordagens distintas operando em paralelo. Os detalhes da arquitetura do mecanismo de detecção de fraude proposto serão apresentados na seção 4.

## 2.7 Métricas de desempenho

As métricas de desempenho, normalmente utilizadas para avaliar o desempenho de detectores de fraudes são as seguintes (KOY, Y., et al., 2004):

**Taxa de verdadeiro positivo (*Tvp*)** ou sensibilidade é a fração de medições fraudulentas que foram corretamente classificadas como fraudulentas.

**Taxa de falso positivo (*Tfp*)** ou taxa de falsos alarmes é a fração de medições legítimas que foram incorretamente classificados como fraudulentas.

**Taxa de verdadeiro negativo (*Tvn*)** é a fração de medições legítimas que foram corretamente classificadas como legítimas.

**Taxa de falso negativo (*Tfn*)** é a fração de medições fraudulentas que foram incorretamente classificadas como sendo legitimas.

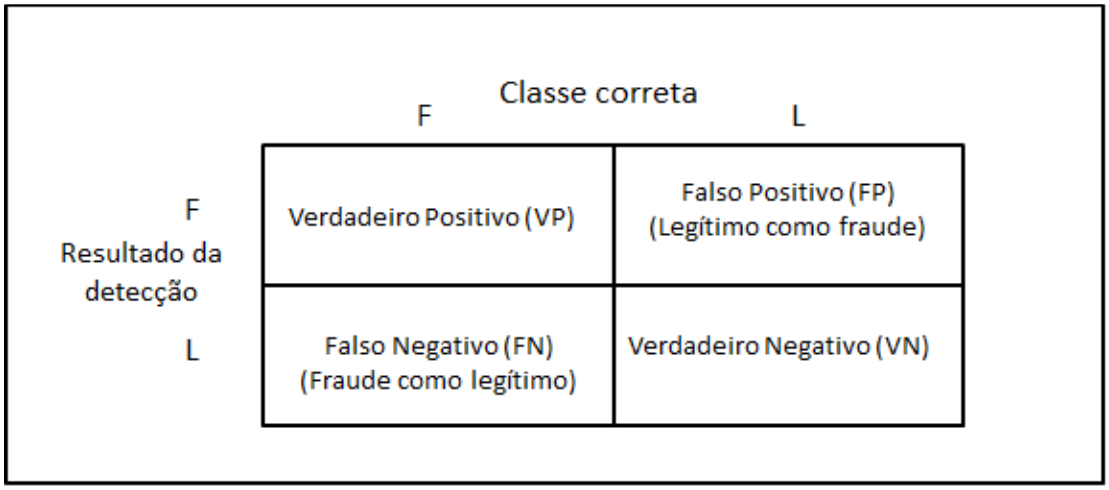
**Exatidão (*Ex*)** é a proporção do número de medições, legítimas e fraudulentas, corretamente identificadas em relação ao número total de medições realizadas.

**Precisão (*Pr*)** é a fração das medições classificadas como fraudulentas que estavam corretas.

As métricas de desempenho podem ser derivadas a partir de uma tabela conhecida como *matriz de confusão* (FAWCETT, T., 2006). A matriz de confusão de uma hipótese X oferece uma medida efetiva do modelo de classificação, ao mostrar o número de classificações corretas *versus* as classificações preditas para cada classe.

Como um detector de fraudes pode ser considerado um classificador de duas classes, P (Positiva ou Fraude) e N (Negativa ou Legítima), o resultado da medição de consumo de energia elétrica de uma unidade consumidora poderia ser classificado de quatro formas diferentes, de acordo com a matriz de confusão representada pela Figura 6.

Figura 6 - Matriz de confusão




Fonte: Kovach (2011)

Onde F corresponde a classe Fraude, L a classe legítima, VP ao número de positivos (fraudes) classificados corretamente, FP ao número de negativos (legítimos) classificados incorretamente, VN ao número de negativos (legítimos) classificados corretamente e FN ao número de positivos (fraudes) classificados incorretamente.

Os componentes da diagonal principal da matriz de confusão fornecem o número de medições corretamente classificadas para cada classe correspondente.

As métricas de desempenho são calculadas da seguinte forma:

Taxa de verdadeiro positivo ou sensibilidade: TVP = VP / P

Taxa de falso positivo: TFP = FP / N

Taxa de verdadeiro negativo: TVN = VN/ N

Taxa de falso negativo: TFN = FN / P

Exatidão: Ex = VP + VN / (P + N)

Precisão: Pr = VP / (VP + FP)

Onde, P = VP + FN corresponde ao número total de medições de consumo de energia irregulares (fraudulentas) e N = VN + FP ao número total de medições de consumo de energia regulares (legítimas).

## 2.8 Métodos de detecção de fraudes

Alguns detectores de fraudes utilizam métodos baseados na comparação do perfil de comportamento observado, com padrões de perfis já conhecidos. Ou seja, o detector de fraude, possui uma base de dados com os perfis de comportamentos fraudulentos mais conhecidos. Caso o perfil de comportamento observado coincida com um dos perfis de comportamento fraudulentos conhecidos, um alerta é disparado.

Outros detectores de fraudes, no entanto, baseiam suas análises em métodos estatísticos, por meio dos quais se busca identificar desvios significativos de comportamento frente ao comportamento padrão e assim, possam indicar suspeita de fraude.

Em qualquer um dos casos, os detectores de fraudes têm por objetivo classificar uma medição ou evento como legítima ou fraudulenta, baseando-se na análise do comportamento observado. Com base nessa análise, os métodos de detecção de fraude, de modo geral, geram valores que são comparados com limiares pré-estabelecidos. Portanto, caso o comportamento observado seja diferente do comportamento característico, ou seja, apresente um valor superior ao limiar pré-estabelecido, o detector de fraudes emite um alerta.

Os métodos de detecção de fraude podem ser classificados como supervisionados e não supervisionados.

Métodos supervisionados são aqueles que utilizam amostras de comportamentos considerados normais e anômalos como parâmetro para comparação e classificação de observações de novos comportamentos. Já os métodos não supervisionados, segundo o mesmo autor, procuram identificar anormalidades com relação aos comportamentos usuais (KOVACH, 2011).

### 2.8.1 Métodos estatísticos

“Métodos estatísticos utilizam métricas e modelos estatísticos para determinar as variações de comportamento dos usuários.” (KOVACH, 2011). Esses métodos baseiam-se nas análises absoluta e diferencial.

Na abordagem baseada em análise absoluta, a detecção de fraude é realizada por meio do estabelecimento de algum critério da comparação, a partir de um ou mais dados de uma medição com valores fixos preestabelecidos, denominados limiares (KOVACH, 2011). No entanto, vale ressaltar que, embora certas mudanças no padrão de consumo de algumas unidades consumidoras possam ser indícios de fraude, em outras elas podem ser consideradas aceitáveis, como, por exemplo, um ligeiro aumento do consumo de energia elétrica em virtude da aquisição de um novo eletrodoméstico. Este fato justifica a necessidade de adoção de outro método, que, atuando conjuntamente com o primeiro, poderia ser

utilizado para fortalecer a possibilidade de ocorrência de verdadeiro positivo. O método em questão será apresentado e detalhado na seção 4 deste trabalho.

Já na abordagem baseada em análise diferencial, o padrão de consumo de energia elétrica de uma unidade consumidora é monitorado e comparado com o seu histórico de consumo de energia.

Como resultado da comparação entre o novo padrão de consumo observado e o padrão de consumo anterior, o mecanismo gera um valor que determina o grau de anormalidade do comportamento observado. Caso não seja identificada qualquer anormalidade, o novo padrão de consumo é incorporado ao padrão anterior visando à adaptação das variações ao padrão de consumo de energia elétrica da unidade consumidora. Se, no entanto, o sistema detectar uma mudança muito significativa do padrão de consumo de energia elétrica dentro de um curto espaço de tempo, um alarme é disparado.

### 2.8.2 Métodos baseados em regras

Métodos baseados em regras podem ser definidos como algoritmos de aprendizado supervisionado que produzem classificadores utilizando regras na forma "... SE ⟨antecedente⟩ ENTÃO ⟨consequente⟩ relacionando dados ou fatos (no antecedente, também denominado de premissa ou condição) a alguma ação ou resultado (no consequente, também denominado de conclusão)..." (ZUBEN, 2011).

Como os sistemas que utilizam métodos baseados em regras são apenas capazes de detectar as fraudes identificadas por essas regras, eles precisam ser constantemente atualizados com novas regras à medida que novas fraudes são descobertas. Além disso, podem cometer erros durante a inferência, devido à dificuldade de lidar com ambiguidades e regras conflitantes e pela necessidade de aprendizado e atualização constante da base de conhecimento.

Finalmente, segundo Kovach (2011), “Uma das consequências dessa abordagem pode ser a necessidade de uma grande quantidade de memória de armazenamento o que pode acarretar atraso no processamento.”.

### 2.8.3 Métodos baseados em redes neurais artificiais

Redes neurais artificiais são sistemas computacionais compostos por várias unidades de processamento, interconectadas por canais de comunicação associados a determinados pesos, que são ajustados à medida que ocorrem interações entre as unidades de processamento da rede, de acordo com os padrões apresentados. O processo de aprendizado decorre justamente da identificação do conjunto apropriado de pesos para que a rede se comporte da forma desejada.

Redes neurais artificiais são utilizadas como técnica de reconhecimento de padrões em detecção de fraudes. Neste caso, durante a fase de aprendizado, os parâmetros da rede são ajustados para associar à saída uma classe (legítima ou fraudulenta), de acordo com um padrão de entrada (KOVACH, 2011).

Se a rede não possuir nenhuma classe associada a uma determinada entrada, então, a rede fornecerá uma saída que corresponda ao padrão que mais se aproximar daquele que foi fornecido na entrada (MORANDI, M.; ZULKERINE, M.; 2004; BEBAR, H. et al., 2002).

Segundo Kovach (2011), uma desvantagem dos métodos baseados em redes neurais artificiais é que a arquitetura da rede e a determinação dos pesos das conexões apenas são definidas após considerável número de tentativas e erros.

## 2.9 Síntese do capítulo

O presente capítulo apresentou uma revisão dos principais conceitos, fundamentos e trabalhos relacionados a *Smart grids* e à detecção de fraudes, dentro da perspectiva abordada pelo trabalho. Neste capítulo discorreu-se os benefícios da adoção de *Smart grids* tanto sob a perspectiva dos consumidores, quanto das concessionárias. Traçou-se um panorama da adoção de Smart grids e apresentou-se as principais características e componentes das redes elétricas inteligentes e levantou-se o problema da detecção de fraudes em consumo de energia elétrica. Finalmente, comentou-se sobre os principais métodos de detecção de fraudes.

No próximo capítulo é apresentada uma revisão bibliográfica sobre confiança computacional e sobre a Teoria Matemática de Evidências de Dempster-Shafer, buscando identificar aspectos importantes dentro da perspectiva da detecção de fraudes.

# 3 TEORIA DE EVIDÊNCIA DE DEMPSTER-SHAFER

Esta seção tem por objetivo apresentar uma revisão conceitual sobre confiança computacional e a utilização da Teoria Matemática de Evidências de Dempster-Shafer na perspectiva da detecção de fraudes em *Smart grids*.

## 3.1 Definição

A Teoria de Evidência de Dempster-Shafer (TDS) é um modelo que permite o tratamento e a representação matemática de incerteza ou ignorância em domínios complexos. Trata-se de uma alternativa à Teoria de Probabilidade. Enquanto na Teoria de Probabilidade a probabilidade deve ser igualmente distribuída, na Teoria de Evidência de Dempster-Shafer é possível atribuir medidas de incerteza a um conjunto de hipóteses disjuntas (KOVACH, 2011).

Na Teoria de Evidência de Dempster-Shafer o conjunto das hipóteses primitivas é chamado de domínio do problema ou quadro de discernimento e é denotado por Θ. O quadro de discernimento é um conjunto de elementos, dentro de um ambiente, que podem ser assumidos como possíveis respostas (KOVACH, 2011).

Um quadro de discernimento é constituído por todos os subconjuntos, formados pela disjunção de seus elementos dentro de um determinado domínio ou ambiente.

A Teoria de Evidência de Dempster-Shafer assume que, para qualquer domínio de problema Θ é exaustivo, no sentido de ser completo (conter toda possível hipótese primitiva) e as hipóteses primitivas em Θ são mutuamente exclusivas (UCHÔA, J., Q., et al., 2000).

Cada subconjunto de Θ, formado pela disjunção de seus elementos, pode ser interpretado como uma possível hipótese, dando origem a $2^{\Theta}$ possíveis hipóteses.

Uma das vantagens da utilização da Teoria de Evidência de Dempster-Shafer é que, a despeito de qualquer conhecimento, *a priori* é possível utilizá-la para detectar fraudes ou comportamentos anômalos não observados, a partir da combinação de evidências obtidas de diversas fontes de detecção. Além disso, a Teoria de Evidência de Dempster-Shafer permite exprimir ignorância a partir da atribuição de medidas de incerteza a um conjunto de hipóteses disjuntas.

### 3.1.1 Função de massa e ignorância

A Teoria de Evidência de Dempster-Shafer utiliza uma função de atribuição de probabilidade básica, chamada de função de massa, simbolizada pela letra *m*. A massa é atribuída aos subconjuntos do ambiente aos quais se deseja atribuir confiança (*belief*).

Considerando-se uma determinada evidência, a função de massa, *m*, atribui a cada possível subconjunto de Θ e, inclusive, ao próprio Θ (ou seja, a 2Θ), um valor no intervalo [0, 1], onde 0 representa desconfiança e 1 representa confiança, de tal formal que a soma de todas essas atribuições, incluindo o valor atribuído ao próprio Θ, seja 1. Por definição, o número 0 deve ser atribuído ao conjunto vazio, uma vez que o conjunto vazio corresponde à hipótese falsa.

Se nenhuma medida de confiança for atribuída a um subconjunto específico, este subconjunto será considerado *sem nenhuma confiança* (*nonbelief*). A medida de confiança não atribuída pela evidência a nenhum subconjunto de Θ, denominada confiança não atribuída *m*(Θ), é atribuída ao próprio ambiente, e não à refutação das hipóteses que receberam confiança. A confiança que reputa uma hipótese é denominada desconfiança (*disbelief*), e, portanto, não tem o mesmo significado que *sem nenhuma confiança*.

Todo subconjunto do quadro de discernimento, ou espaço amostral, cuja massa de atribuição de confiança for diferente de 0, ou seja, todo subconjunto para o qual m > 0, é chamado de elemento focal. (SILVA, J. D. S, 2000)

Supondo que, a título de ilustração, em um domínio Θ = {Fraude, ¬Fraude}, um detector de fraudes indicasse uma confiança de 0,6 na evidência de que os dados de consumo de energia elétrica reportados pelo medidor de uma unidade consumidora fossem incompatíveis com os dados reais, ou seja, fraude, a atribuição de massa para o conjunto {Fraude}, seria expressa por m ({ Fraude }) =0,6. Neste exemplo, o resto da confiança, ou seja, 0,4 (1 – 0,6) seria deixado para o ambiente Θ.

Enquanto na Teoria de Probabilidade o valor deixado para o ambiente Θ representa o grau de desconfiança na fraude, na Teoria de Evidência de Dempster-Shafer isto significa que se confia que a medição dos dados de consumo seja fraude com um grau de 0,6 e reserva-se um parecer de 0,4 tanto na desconfiança como na confiança adicional de que realmente seja uma fraude.

### 3.1.2 Regra de combinação de Dempster-Shafer

A regra de combinação da Teoria de Evidência de Dempster-Shafer tem por objetivo combinar evidências independentes, obtidas por meio de diversas fontes de observação.

Duas massas de evidências, $m_1(Z)$ e $m_2(Z)$, obtidas por meio de duas fontes independentes, sobre o mesmo quadro de discernimento, utilizando-se a regra de combinação de Dempster-Shafer poderiam ser combinadas da seguinte forma:

$$m_3(Z) = m_1(Z) \oplus m_2(Z) = \frac{\sum_{X \cap Y = Z} m1(X).m2(Y)}{1 - K} \qquad (1)$$

Considerando o exemplo anterior e supondo que um segundo detector de fraude apontasse, na mesma medição de dados de consumo de energia elétrica daquela unidade consumidora, uma evidência de fraude da ordem de 0,8, as massas de confiança dos dois detectores de fraudes, portanto, seriam:

$m_1$ ({ Fraude} ) = 0,6 e $m_1$ (Θ ) = 0,4

$m_2$ ({ Fraude} ) = 0,8 e $m_2$ (Θ ) = 0,2

Pela regra de combinação de Dempster, as evidências seriam então combinadas, conforme apresentado na Tabela 1.

Tabela 1 - Tabela de intersecções e produtos de $m_1$ e $m_2$

| | $m_2$ ({Fraude} ) = 0,8 | $m_2$ (Θ) = 0,2 |
|---|---|---|
| $m_1$ ({ Fraude} ) = 0,6 | {Fraude} = 0,48 | {Fraude} = 0,12 |
| $m_1$ (Θ ) = 0,4 | {Fraude} = 0,32 | Θ = 0,08 |

Fonte: elaborado pelo autor

De acordo com a regra de combinação de Dempster, os conjuntos comuns resultantes seriam somados:

$m_3$ ({ Fraude} ) = $m_1 \oplus m_2$ ({ Fraude} ) = 0,48 + 0,32 + 0,12 = 0,92

$m_3$ (Θ) = $m_1 \oplus m_2$ (Θ) = 0,08

onde, $m_3$ ({ Fraude} ) representa a confiança na evidência combinada de fraude e $m_3$ (Θ), um valor adicional, que poderia ser somado à confiança de 0,92 no conjunto {Fraude} para reforçar a evidência de fraude.

Dessa forma, o intervalo de confiança da evidência de fraude variaria entre 0,92 e 1,0 e seria representado por [0,92 e 1.0], no qual, o limite inferior é conhecido como Confiança (*Bel – Belief*) e o limite superior, como Potencial de Confiança ou Confiança Plausível (*Pls – Plausibility*).

Neste exemplo, o valor de K da função de combinação é igual a zero, pois as intersecções não resultaram em nenhum conjunto vazio.

### 3.1.3 Função de confiança

A função de confiança (Bel) corresponde à soma de todas as atribuições de massas de confiança.

Segundo Kovach (2011), a função de confiança (*Bel*) fornece a confiança total de um conjunto e de todos os seus subconjuntos, ou seja, *Bel* corresponde "... à soma de todas as massas que dão suporte a um conjunto."

Como as funções de confiança podem ser definidas em termos de massa, a combinação de duas funções de confiança pode ser expressa em termos de soma ortogonal das massas de um conjunto e de todos os seus subconjuntos (KOVACH, 2011). Por exemplo:

$Bel_1 \oplus Bel_2 (\{A,B\}) = m_1 \oplus m_2 (\{A,B\}) + m_1 \oplus m_2 (\{A\}) + m_1 \oplus m_2 (\{B\})$

### 3.1.4 Intervalo de confiança (IC)

Bel (Y) exprime o grau com que a evidência suporta a hipótese Y, "... isto é, fornece um limite inferior de confiança." (KOVACH, 2011). Já Bel (Y´) exprime o grau com que a hipótese Y é refutada.

Pls (Y) = 1 – Bel (Y´) representa a confiança total não atribuída a Y', estabelecendo um limite superior de confiança a Y.

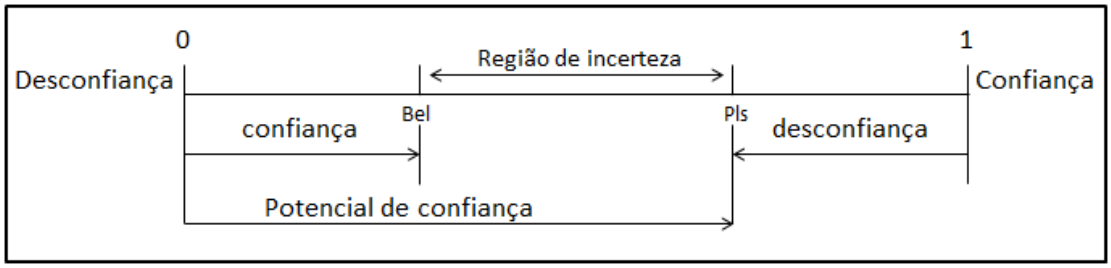
A diferença entre Pls (Y) e Bel (Y) representa o grau de incerteza ou ignorância em relação a Y.

Portanto, o intervalo de confiança de um conjunto Y é definido por:

$$IC (Y) = [ Bel (Y), Pls (Y) ] \qquad (2)$$

A Figura 7 ilustra as relações entre as grandezas de um intervalo de confiança.

Figura 7 - Intervalo de confiança




Fonte: Kovach (2011)

Supondo agora, que no exemplo anterior, um terceiro detector de fraude, apresentasse uma evidência de 0,75 de que os dados de consumo de energia elétrica reportados pelo medidor da unidade consumidora fossem, na verdade, condizentes com os dados reais, ou seja, uma evidência conflitante da ocorrência de fraude, Neste caso, considerando, a confiança na evidência de fraude resultante da combinação das massas de confiança dos dois detectores de fraudes anteriores, conforme a Tabela 2, $m_1 \oplus m_2$ ({ Fraude} ) = 0,48 + 0,32 + 0,12 = 0,92 e $m_1 \oplus m_2$ (Θ) = 0,08, têm-se agora:

$m_3$ ({¬Fraude}) = 0,75 e $m_3$ (Θ) = 0,25

Tabela 2 - Tabela de intersecções e produtos entre $m_1$, $m_2$ e $m_3$

| | $m_1 \oplus m_2$ ( {Fraude} ) = 0,92 | $m_1 \oplus m_2$ (Θ) = 0,08 |
|---|---|---|
| $m_3$ ({¬Fraude} ) = 0,75 | { Ø } 0,69 | { ¬Fraude } 0,06 |
| $m_3$ (Θ) = 0,25 | { Fraude } = 0,23 | Θ 0,02 |

Fonte: elaborado pelo autor

Neste caso, o conjunto vazio Ø ocorre porque {Fraude} e {¬Fraude} não tem nenhum elemento em comum. O fator K, igual à soma das massas dos conjuntos vazios resultantes da intersecção é igual a 0,69. Portanto, o fator de normalização é 1 – K = 1 – 0,69 = 0,31.

Aplicando a regra de combinação de Dempster, têm-se:

$m_1 \oplus m_2 \oplus m_3$ ( {¬Fraude} ) = 0,06/0,31 = 0,194

$m_1 \oplus m_2 \oplus m_3$ ( {Fraude} ) = 0,23/0,31 = 0,742

$m_1 \oplus m_2 \oplus m_3$ (Θ) = 0,02/0,31 = 0,065

A confiança total no subconjunto {Fraude} será, portanto:

Bel ( {Fraude} ) = $m_1 \oplus m_2 \oplus m_3$ ( {Fraude} ) = 0,742, e

Bel ( {¬Fraude} ) = $m_1 \oplus m_2 \oplus m_3$ ( {¬Fraude} ) = 0,194

Pls ( {Fraude} ) = 1 – Bel ({¬Fraude } ) = 1- 0,194 = 0,806

Portanto, o novo intervalo de confiança será:

IC ( {Fraude} ) = [0,742, 0,806 ]

O suporte a hipótese (Bel) e o potencial de confiança (Pls) para {Fraude} foram reduzidos pela evidência conflitante de {¬Fraude}.

## 3.2 Síntese do capítulo

Neste capítulo discorreu-se uma definição conceitual sobre confiança, tanto sob o aspecto humano quanto sob o aspecto computacional. Apresentou-se a Teoria Matemática de Dempster-Shafer, objeto principal deste trabalho. Com base no entendimento da regra de combinação de Dempster-Shafer, concluiu-se que duas evidências, obtidas por meio de duas fontes independentes dentro do mesmo quadro de discernimento podem ser combinadas a fim de gerar-se uma evidência final. Este aspecto será particularmente importante dentro da perspectiva do mecanismo de fraude que será proposto no próximo capítulo.

# 4 PROPOSTA DE DETECÇÃO DE FRAUDES DE CONSUMO DE ENERGIA EM *SMART GRIDS* UTILIZANDO A TEORIA DE DEMPSTER-SHAFER

Este capítulo apresenta a proposta de mecanismo de detecção de fraude de consumo de energia elétrica utilizando a teoria de evidência de Dempster-Shafer.

## 4.1 Considerações iniciais

O mecanismo de detecção de fraude proposto neste trabalho baseia-se no método de detecção de fraudes em transações financeiras via internet em tempo real desenvolvido por Kovach (2011) – adaptação.

A arquitetura do mecanismo de detecção de fraudes proposto neste trabalho será constituída por dois métodos (duas abordagens) operando em paralelo, utilizando, no entanto, classes de atributos diferentes em função das especificidades do domínio de medições de consumo de energia elétrica.

Neste trabalho, diferentemente do trabalho desenvolvido por Kovach (2011) optou-se por utilizar como modelos estatísticos para o cálculo da distância probabilística entre os perfis de consumo atual e histórico, a média móvel simples e ponderada.

O método a ser utilizado para validar a efetividade do mecanismo de detecção de fraudes proposto seguirá os principais itens que sintetizam o objetivo do trabalho:

- determinação do atributo local e global mais adequados para determinar o perfil de consumo das unidades consumidoras com base numa análise empírica dos dados;

- escolha e aplicação dos modelos estatísticos mais adequados para caracterizar evidências de fraude no consumo de energia elétrica com base nos atributos escolhidos e;
- definição de um método para combinar as evidências de fraude e determinar um escore de suspeita de fraude.

## 4.2 Atributos de medição de consumo de energia elétrica

Para fins de detecção de fraudes, uma medição de consumo de energia elétrica é caracterizada por um conjunto de atributos. A determinação dos atributos de medição de consumo de energia elétrica mais apropriados para detecção de fraudes é determinante na capacidade de discriminação do detector, pois é por meio deles que os perfis de atividades individuais são definidos.

### 4.2.1 Perfil de atividade individual

O perfil de atividade individual descreve os aspectos observáveis do comportamento das unidades consumidoras e são definidos a partir dos atributos de medição do consumo de energia elétrica de uma unidade. Eles serão utilizados para se distinguir um comportamento normal de um comportamento anômalo.

Uma métrica representa uma medida quantitativa de uma variável aleatória, acumulada durante um período de medições. As métricas são normalmente obtidas por meio de contagens, temporizações e medições de recursos. Já os modelos estatísticos têm por objetivo determinar se as novas medições serão consideradas normais ou anormais, em relação às medições realizadas anteriormente. Dentre os modelos estatísticos mais comumente utilizados em detecção de fraudes, estão os modelos baseados em média, desvio padrão e limiar fixo.

São exemplos de métricas e modelos estatísticos geralmente utilizados para determinação de perfis de atividades individuais das unidades consumidoras:

- consumo de energia elétrica:
  métrica: medição do consumo de energia elétrica mensal.
  modelo estatístico: média e desvio padrão.
- Demanda máxima e demanda mínima:
  métrica: consumo máximo e mínimo de energia elétrica registrado pela unidade consumidora ao longo dos últimos doze meses.
  modelo estatístico: limiar fixo.
- Demanda por período:
  métrica: medição da demanda de consumo de energia elétrica por período.
  modelo estatístico: média e desvio padrão.

### 4.2.2 Perfil de comportamento

O perfil de comportamento (PC) de uma unidade consumidora será obtido por meio de um vetor de variáveis aleatórias, no qual, cada variável corresponderá a um perfil de atividade individual. O perfil de comportamento de uma unidade consumidora será descrito pela seguinte expressão:

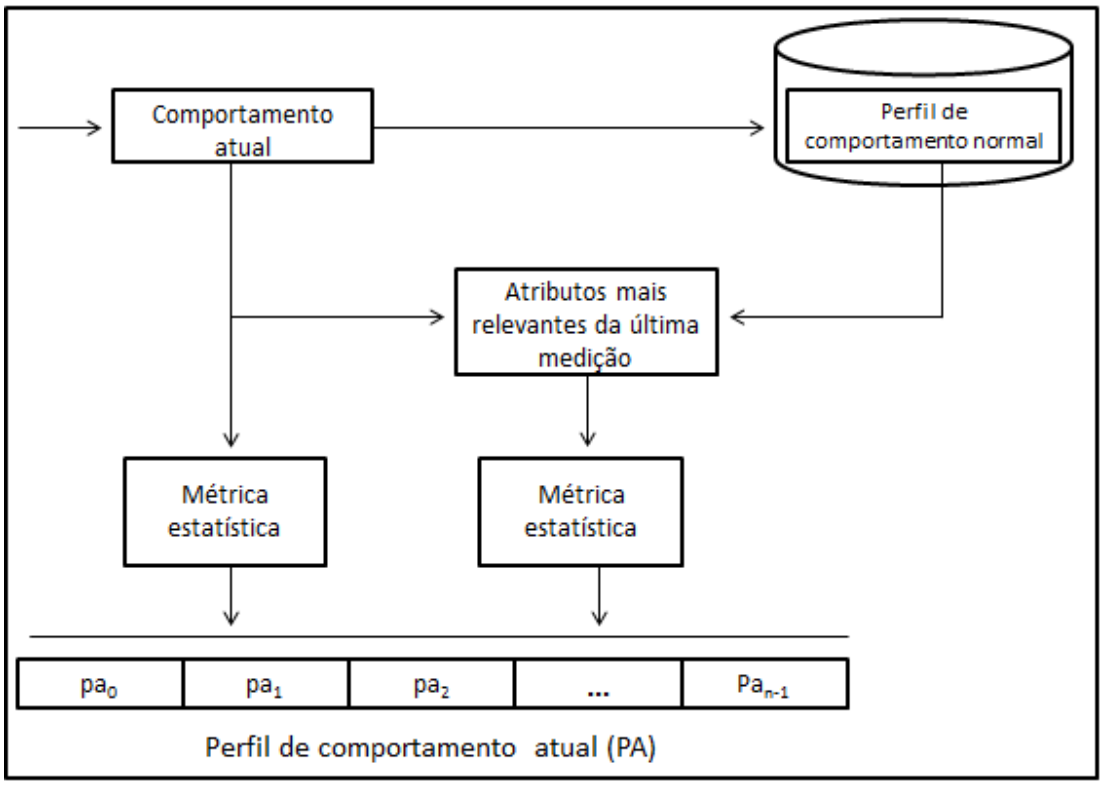
$$PC = \{ pa_0, pa_1, ..., pa_{n-1} \} \qquad (3)$$

Onde, $pa_i$ corresponderá ao perfil de atividade individual do elemento i do vetor.

O objetivo da determinação do perfil de comportamento de cada unidade consumidora é obter um vetor capaz de exprimir valores que representam o comportamento normal esperado de uma unidade consumidora, com base no seu histórico de comportamentos acumulados. Assim, o perfil de comportamento atual de uma unidade consumidora poderá ser comparado com o seu perfil de comportamento normal (considerado ideal) a fim de se verificar se existe alguma inconsistência significativa que possa indicar suspeita de fraude.

A determinação do perfil do comportamento atual de uma unidade consumidora, bem como a verificação de sua consistência com o perfil histórico de comportamento, será realizada pelo módulo de análise diferencial. A Figura 8 ilustra a construção do perfil de comportamento atual de uma unidade consumidora.

Figura 8 - Construção do perfil de comportamento atual




Fonte: adaptado de Kovach (2011)

Dentro da perspectiva apresentada e com base nas características específicas do domínio de aplicação deste trabalho, foram escolhidas empiricamente duas classes de atributos:

- uma centrada no perfil de comportamento das unidades consumidoras em geral; e

- outra centrada no perfil de comportamento das unidades consumidoras irregulares.

Com base nesta classificação, o atributo centrado nas unidades consumidoras em geral, será denominado de atributo local. Já, o atributo centrado no perfil de das unidades consumidoras irregulares será denominado de atributo global.

## 4.3 Características de evidências de fraudes em medições de energia elétrica

Segundo os especialistas consultados, as principais características de evidências de fraudes em medições de consumo de energia elétrica são:

- variação significativa do consumo de energia elétrica;
- queda acentuada do registro de consumo de energia elétrica entre dois meses consecutivos (degrau de consumo);
- pouca diferença de consumo de energia elétrica durante as estações de verão e inverno;
- registro de consumo de energia elétrica nulo;
- histórico de fraudes;
- faturamento pela taxa mínima;
- histórico de inspeção *in loco* na unidade consumidora;
- denúncia;
- menor consumo de energia elétrica que o usual dentro do horário de pico; e
- maior consumo de energia elétrica que o usual durante o horário de tarifa branca.

O mecanismo de detecção de fraudes proposto neste trabalho leva em consideração essas características na determinação dos atributos utilizados na identificação de comportamento anômalo e no levantamento de evidência de indício de fraude.

## 4.4 Descrição geral da arquitetura do mecanismo de detecção de fraudes proposto

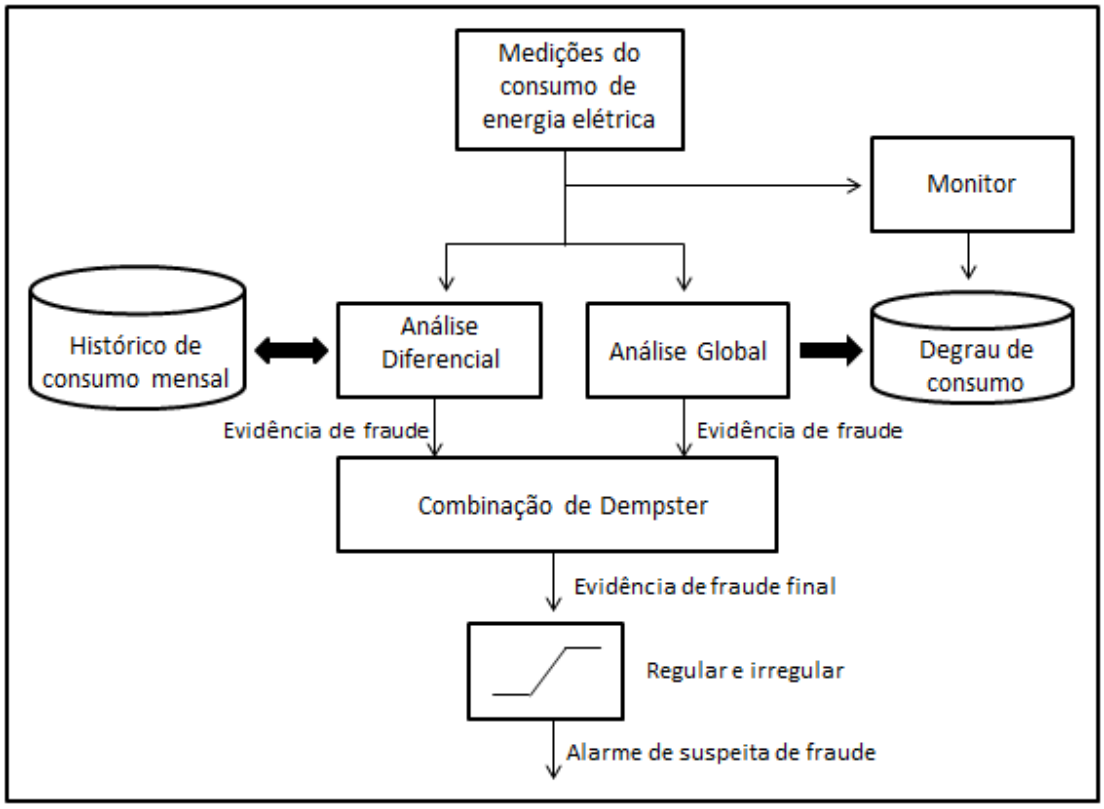
O mecanismo de detecção de fraude proposto neste trabalho baseia-se no método de detecção de fraudes em transações financeiras via internet em tempo real desenvolvido por Kovach (2011). Assim como no trabalho desenvolvido por Kovach (2011), a arquitetura do mecanismo de detecção de fraudes proposto neste trabalho é baseada na utilização de dois métodos (duas abordagens) operando em paralelo, no entanto, utilizando classes de atributos diferentes em função das especificidades do domínio. Além disso, neste trabalho, diferentemente do trabalho desenvolvido por Kovach (2011) optou-se por utilizar a média móvel simples e ponderada como modelos estatísticos para o cálculo da distância probabilística entre os perfis de consumo atual e histórico. A razão desta escolha será detalhada na seção 5.4.1.

As duas abordagens operando em paralelo serão:

- uma baseada em análise diferencial, centrada no consumo de energia elétrica (kWh) da unidade consumidora e que detecta mudanças significativas nos padrões de consumo no decorrer do tempo; e
- outra baseada em análise global, centrada na observação do comportamento global das unidades consumidoras, utilizando como parâmetro, o degrau de consumo.

As evidências de fraude determinadas por meio pelas abordagens apresentadas serão combinadas a fim de se exprimir um valor de suspeita de fraude que, se estiver acima ou abaixo de um limiar preestabelecido, poderá, ou não, disparar um alarme de suspeita de fraude. A Figura 9 apresenta uma visão geral da arquitetura do mecanismo proposto.

Figura 9 - Visão geral da arquitetura do mecanismo proposto




Fonte: adaptado de Kovach (2011)

A arquitetura do mecanismo de detecção de fraudes proposto será baseada em dois pressupostos:

1) a probabilidade de fraude cresce à medida que ocorre variação significativa na medição do consumo de energia elétrica junto à unidade consumidora; e
2) a única forma de se ter certeza de que uma fraude foi perpetrada ocorre quando ela for confirmada pela própria concessionária por meio de uma inspeção.

Os detalhes dos principais aspectos relacionados à arquitetura do mecanismo proposto estão descritos nas próximas seções.

## 4.5 Análise diferencial

Na abordagem utilizando análise diferencial, o perfil de comportamento referente à medição dos dados de consumo atual será comparado com o perfil de comportamento que caracteriza o padrão de consumo de energia elétrica da unidade.

Portanto, para realizar a análise diferencial proposta pelo mecanismo de detecção de fraudes deste trabalho, será necessário dispor de informações sobre o comportamento histórico das medições do consumo de energia elétrica da unidade consumidora. Além disso, será necessário obter-se uma amostra da medição mais recente para que o perfil de comportamento referente ao padrão de consumo atual de uma unidade seja comparado com perfil de comportamento normal dessa unidade. Caso o resultado dessa comparação divirja, significativamente, do padrão de consumo médio, se caracterizará o indício de fraude.

Para realizar esta análise diferencial, o mecanismo proposto utilizará dois perfis: um que descreverá o comportamento da medição de dados de consumo sendo realizada (PA - Perfil Atual); e outro que apresentará o perfil médio histórico de consumo da unidade consumidora (PH – Perfil Histórico).

O Perfil de comportamento da medição atual será calculado por meio das informações fornecidas pelos atributos de medição do consumo de energia elétrica da unidade. O resultado obtido será utilizado para determinar a distância probabilística entre o Perfil Atual e o Perfil Histórico. Esta distância será calculada por meio da aplicação de modelos estatísticos em cada um dos elementos do vetor (perfis individuais de atividade), que por sua vez, irão gerar valores individuais de suspeita de fraude em relação a cada um deles. A união entre todas as distâncias individuais fornecerá o valor da distância total entre o Perfil Atual e o Perfil Histórico.

Uma das maneiras de se obter a distância entre dois perfis individuais é por meio da soma ponderada das distâncias individuais de cada elemento (FERREIRA et al., 2006).

Para verificar se a escolha do modelo estatístico exerce alguma influência significativa no resultado da análise diferencial, o mecanismo de detecção de fraudes proposto neste trabalho utilizará utilizados dois modelos estatísticos diferentes para determinar a distância total entre os dois perfis, PA e PH. Serão utilizados os modelos de média simples e média ponderada.

O valor calculado das distâncias entre os perfis PA e PH - o *dist* (PA, PH) será combinado com as evidências fornecidas pelo módulo de análise global, para produzir um valor total de suspeita de fraude no módulo de combinação de Dempster-Shafer. Se esse valor obtido não significar fraude, o Perfil Histórico (PH) será atualizado com o perfil da medição recente (PA).

### 4.5.1 Atualização do perfil histórico

Como os perfis dos consumidores são dinâmicos, seja pela mudança de comportamento natural do cliente, seja pela alteração da composição familiar (residenciais) ou econômica (comerciais e industriais), essa dinâmica demanda uma reavaliação periódica dos parâmetros adotados pelo sistema.

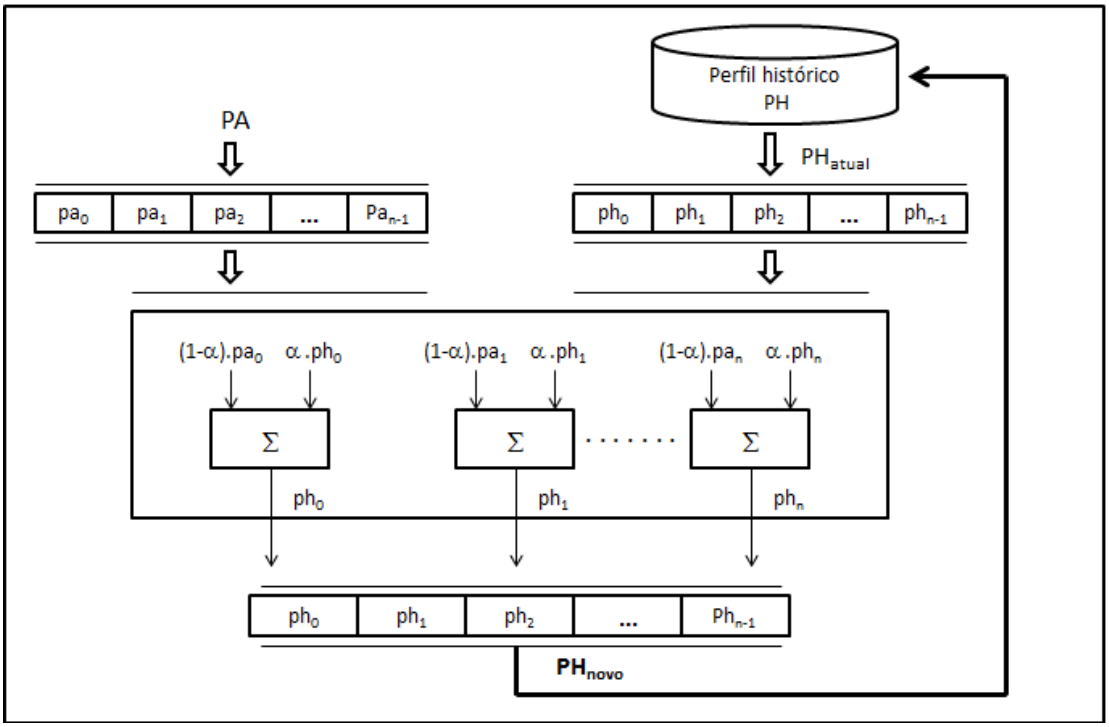
Com o objetivo de permitir que o perfil de comportamento histórico da unidade consumidora se adapte às variações, ele será atualizado, lentamente, para incorporar o perfil das medições de consumo de energia elétrica mais recente. Esta atualização será feita por meio de uma função de média ponderada que determinará a taxa com que os valores antigos serão descartados a cada reavaliação do perfil histórico.

O valor de PH será atualizado com o valor de PA, decaindo os valores de cada elemento de PH por um fator alfa (0 < α< igual 1), por meio da seguinte expressão:

$$ph_i = \alpha . ph_i + (1 - \alpha).pa_i \qquad (4)$$

onde, $Ph_i$ e $pa_i$ representam o i-ésimo elemento de PH e PA, respectivamente; e o valor de α determinará a taxa com que os valores antigos se tornarão irrelevantes, sendo determinados por especialistas. A Figura 10 ilustra o processo de atualização de PH com os valores de PA.

Figura 10 - Atualização de PH




Fonte: Kovach (2011)

## 4.5.2 Inicialização do perfil histórico

A fim de reduzir o número de falsos alarmes causados por suspeita de fraude no consumo de energia elétrica, em virtude da falta de dados sobre o perfil de consumo de novas unidades consumidoras na rede, é necessário escolher o modelo estatístico ou o perfil inicial mais adequado para tratar o problema da inicialização do perfil histórico.

Um modelo estatístico que poderia ser utilizado seria um que fizesse uso do cálculo da média e do desvio padrão, de tal forma que o intervalo de confiança fosse maior no início e diminuísse à medida que fossem coletados mais dados sobre o comportamento da nova unidade consumidora. No entanto, embora essa abordagem pudesse reduzir o número de falsos alarmes causados pela evidência de fraude dos dados de consumo reportados, ela não protegeria o sistema contra comportamentos suspeitos ou mesmo contra unidades consumidoras cujo comportamento fosse, inicialmente, anormal (KOVACH, 2011).

Kovach (2011) apud Cortes e Pregibon (2001) sugere a criação de classes de equivalência de contas, de modo que quando uma nova conta fosse criada, os atributos de comportamento da nova unidade consumidora seriam utilizados para mapear o perfil de consumo inicial dela, para então associá-la a uma classe de equivalência. Ainda de acordo com os mesmos autores, a inicialização do perfil histórico não precisaria ser exata, pois, desde que um perfil inicial tenha sido estabelecido, o perfil histórico seria atualizado com os dados individuais reais.

Ferreira (2006) propõe como solução para o problema da inicialização do perfil histórico a abertura de uma pequena janela de tempo no início das atividades, durante a qual não se detectaria fraudes.

O presente mecanismo de detecção de fraudes de consumo de energia elétrica proposto propõe como solução para o problema da inicialização do perfil histórico das novas unidades consumidoras a adoção de classe de consumo equivalente à da nova unidade consumidora e a utilização de um modelo estatístico baseado em média móvel ponderada, de modo que, à medida que ocorressem novas medições, as medições antigas seriam descartadas e as novas médias seriam calculadas considerando-se apenas os n períodos mais recentes.

## 4.6 Análise global

O objetivo do módulo de análise global será fortalecer as evidências de fraude determinadas pelo módulo de análise diferencial.

Como atributo global, será utilizado o degrau de consumo em decorrência do fato de que estatisticamente unidades consumidoras irregulares apresentam degrau de consumo. Além de se tratar um atributo comumente utilizado pelas concessionárias de energia para selecionar unidades consumidoras suspeitas de fraude para inspeção, a escolha deste atributo deu-se com base numa análise feita na base de dados por meio da qual, constatou-se empiricamente que de fato, boa parte das unidades consumidoras inspecionadas e identificadas a priori como sendo ou irregulares apresentavam degrau de consumo - queda acentuada no consumo de energia elétrica verificada entre dois meses consecutivos.

## 4.7 Combinação de evidências

A combinação dos resultados de vários detectores independentes pode apresentar um desempenho melhor do que o resultado obtido por meio da utilização de um único detector (KOVACH, 2011 apud SINGH, R. et al., 2006; CHEN, Q; AICKELIN U., 2006).

No mecanismo proposto, as evidências de fraude, determinadas pelos módulos individuais de detecção de fraudes, serão combinadas por meio da utilização da teoria matemática de Dempster-Shafer (TDS) visando inferir e exprimir um valor que represente a confiança da medição dos dados de consumo de energia elétrica das unidades consumidoras.

O primeiro passo para se utilizar a regra da combinação da Teoria de Evidência de Dempster-Shafer é mapear as evidências de fraude geradas pelos módulos de detecção em número probabilístico, m(*f*) (KOVACH, 2011).

No mecanismo proposto, as evidências de fraudes determinadas pelos dois módulos já serão fornecidas em valores probabilísticos, não existindo, portanto, a necessidade de se fazer qualquer mapeamento.

A regra da combinação de Dempster-Shafer fornecerá uma função para calcular o valor total de duas evidências. Dadas as massas de duas evidências de fraude $m_1(f)$ e $m_2(f)$, elas poderão ser combinadas em uma terceira massa $m_3(f)$ pela seguinte expressão (KOVACH, 2011):

$$m_3(f) = m_1(f) \oplus m_2(f) = \frac{\sum x \cap y = f \; m1(x).m2(y)}{1-K} \quad (5)$$

O quadro de discernimento Θ no domínio de detecção de fraudes é constituído de dois valores mutualmente exclusivos, isto é,

$$\Theta = \{ f, -f \}$$

Onde, *f* = fraude; e *–f* = legítimo.

O conjunto de todas as hipóteses possíveis de Θ corresponde a todos os subconjuntos de Θ e incluindo ele mesmo. Este conjunto, denotado por $2^{\Theta}$ é constituído de três possíveis hipóteses, {*f*}, {*-f*} e Θ = { *f*, *-f* } (representando a incerteza) (KOVACH, 2011).

Supondo que o módulo tem uma evidência de fraude com probabilidade α, as massas a serem assinaladas seriam:

$m(f) = \alpha$

$m(-f) = 0$

$m(\Theta) = 1- \alpha$

Com base nisso, o resultado da combinação de Dempster, $m_3(f) = m_1(f) \oplus m_2(f)$, será reduzidas a:

$m_3(f) = m_1(f).\, m_2(f) + m_1(f).\, m_2(\Theta) + m_1(\Theta).\, m_2(f)$

Supondo que um detector obtenha uma evidência de fraude com $m_1(f) = 0,2$ e outro detector, uma evidência de fraude com $m_2(f)=0,4$

Neste caso,

$m_1(-f) =0$ e $m_1(\Theta) = 0,8$; e

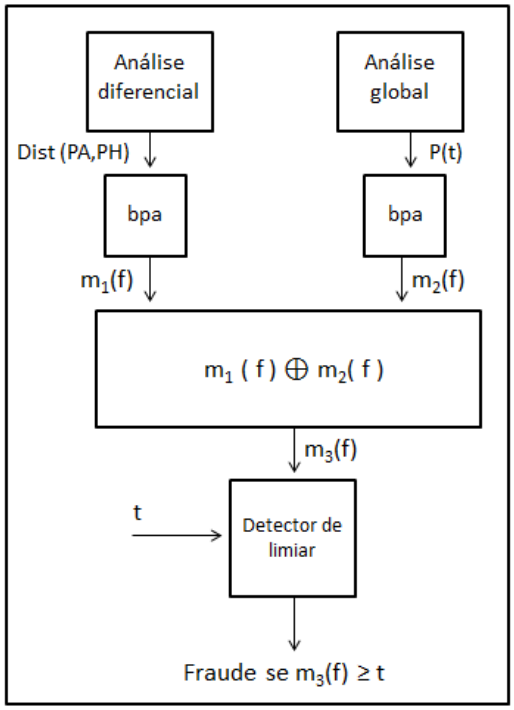
$m_2(-f)= 0$ e $m_2(\Theta) = 0,6$; e

$m_3(f)= 0,2 . 0,4 + 0,2. 0,6 + 0,8. 0,4 = 0,08 + 0,12 + 0,32 = 0,52$

O resultado final será obtido por meio da aplicação de um limiar *t* ao $m_3(f)$, de tal forma que, se $m_3(f) \geq t$, fraude, caso contrário, legítimo.

A Figura 11 ilustra em blocos o módulo de combinação do sistema proposto utilizando a Teoria de Evidência de Dempster-Shafer.

Figura 11 - Módulo de combinação de evidências




Fonte: Kovach (2011)

Vale ressaltar que a função de combinação de Dempster é apresentada para duas massas de evidências somente. Caso haja mais evidências a serem

combinadas, recomenda-se que a combinação de evidências seja feita de duas em duas. Ou seja, combinam-se duas evidências, para em seguida combinar o resultado obtido dessa combinação com uma terceira evidência e assim por diante. Caso um novo módulo de detecção seja adicionado ao sistema existente, o ideal seria que ele já fornecesse a evidência de fraude expressa em valores probabilísticos (KOVACH, 2011).

## 4.8 Síntese do capítulo

Neste capítulo discorreu-se sobre as principais características de evidências de fraudes em medições de energia elétrica.

Analisaram-se os principais atributos de medição de consumo de energia elétrica e definiram-se as classes de atributos mais adequados para se determinar o perfil de consumo das unidades consumidoras. Especificamente, decidiu-se por usar uma classe centrada no perfil de comportamento das unidades consumidoras em geral e outra centrada no perfil de comportamento das unidades consumidoras irregulares. Com base na definição destas duas classes de atributos, escolheram-se os modelos estatísticos média móvel simples e ponderada para caracterizar evidências de fraude no consumo de energia elétrica.

Apresentou-se os detalhes da implementação da arquitetura do mecanismo de detecção de fraudes proposto, em especial os módulos de análise diferencial e análise global.

Por fim, elegeu-se um método para combinar as evidências de fraude determinadas pelos módulos individuais (local e global) e determinar um escore de suspeita de fraude, no caso, a regra da combinação da Teoria Matemática de Dempster-Shafer.

No próximo capítulo são analisados os resultados decorrentes da implementação do método para validação desta proposta.

# 5 VALIDAÇÃO DA PROPOSTA

Este capítulo tem por objetivo realizar a validação do mecanismo de detecção de fraudes proposto no capítulo 4.

## 5.1 Considerações iniciais

O objetivo da validação do mecanismo de detecção de fraudes proposto foi mostrar que a utilização de duas abordagens - uma centrada no perfil de comportamento das unidades consumidoras em geral (local) e outra centrada no perfil de comportamento das unidades consumidoras irregulares (global) – quando combinadas, melhora a capacidade de discriminação do detector de fraude em relação à utilização de uma única abordagem.

A análise local, realizada com base na análise diferencial do registro de consumo mensal (kWh) de cada unidade consumidora, teve por objetivo determinar se o padrão de consumo de energia atual diferia muito da média histórica da unidade consumidora.

Para fortalecer as evidências de fraude determinadas pelo módulo de análise diferencial e validar a hipótese inicial do trabalho de que o uso de um atributo local em conjunto com um atributo global aumenta a taxa de especificidade, foi utilizado o módulo de análise global baseado na variação brusca de consumo de energia entre 2 meses consecutivos.

Finalmente, os valores das evidências determinadas pelos módulos de análise local e global foram então, combinados utilizando-se a Teoria Matemática de Dempster-Shafer a fim de gerar-se um escore final de suspeita de fraude.

A validação do desempenho do mecanismo foi feita por meio da utilização das curvas ROC e AUC.

Por fim, foi realizada uma análise crítica dos resultados decorrentes da validação do mecanismo proposto.

## 5.2 Método de validação da proposta

O método utilizado para validar o mecanismo de detecção de fraude baseou-se, essencialmente, em:

- escolha dos atributos locais e globais para determinar o perfil de consumo das unidades consumidoras, com base em uma análise empírica dos dados;
- escolha e aplicação dos modelos estatísticos para caracterizar as evidências de fraude no consumo de energia elétrica, com base nos atributos escolhidos;
- combinação das evidências obtidas, por meio dos módulos de análise, local e global utilizando a Teoria Matemática de Dempster-Shafer e;
- avaliação do desempenho do mecanismo de detecção de fraudes de consumo de energia elétrica.

## 5.3 Dados para validação

A validação da efetividade do mecanismo de detecção de fraudes proposto foi realizada por meio de simulações, utilizando uma amostragem estratificada gerada a partir de uma base de dados fornecida por uma concessionária de energia elétrica, contendo dados reais de medições de consumo de energia elétrica de unidades consumidoras de um município do interior de São Paulo. Algumas unidades consumidoras ainda não haviam sido inspecionadas enquanto outras, já haviam sido inspecionadas e identificadas a priori como sendo regulares ou irregulares.

Originalmente, a base de dados fornecida pela concessionária continha dados de consumo de 42.396 unidades consumidoras. Desse total, foram desconsideradas as unidades consumidoras que não haviam sido inspecionadas.

Durante a fase de pré-processamento dos dados, após a limpeza dos registros incompletos, inconsistentes e dos ruídos, o número total de unidades consumidoras foi reduzida para 13.625. Dessas, 12.375 (90,8%) haviam sido classificadas como regulares e 1250 (9,2%) classificadas como irregulares. A Tabela 3 mostra a quantidade de unidades consumidoras por estrato.

Tabela 3 - Quantidade de unidades consumidoras por estrato

| **CLASSE/FASE** | **REGULARES** | **IRREGULARES** | **TOTAL** |
|---|---|---|---|
| RESIDENCIAL | 11.669 | 1212 | 12.881 |
| COMERCIAL | 551 | 31 | 582 |
| INDUSTRIAL | 48 | 1 | 49 |
| OUTROS | 107 | 6 | 113 |
| TOTAL | 12.375 | 1250 | 13.625 |

Fonte: elaborado pelo autor

Para que a massa de dados original fosse convertida em uma base de dados menor, sem, no entanto, perder as propriedades e a representatividade dos dados originais, foi empregada a técnica de amostragem aleatória estratificada. A técnica de amostragem proporcional estratificada consiste na separação dos dados em diversos estratos, tomando-se de cada estrato uma amostra proporcional à representativa do estrato em relação à população. Uma amostra fornecerá resultados de qualidade semelhantes aqueles produzidos pelo conjunto de dados completos se ela for representativa. Uma amostra é representativa se ela mantiver as mesmas propriedades (de interesse) do conjunto de dados original.

Para a determinação do erro e do tamanho da amostra (n) aplica-se a seguinte formulação:

$$erro = Z_{\propto} * \sqrt{\frac{p*(1-p)}{n}} \quad (6)$$

$$n = \left(\frac{Z_{\propto} * \sqrt{p*(1-p)}}{erro}\right)^2 \quad (7)$$

A fim de se reduzir o tamanho do conjunto de dados para facilitar a aplicação do mecanismo proposto, buscou-se obter uma amostra cujo tamanho, embora fosse reduzido, assegurasse um erro amostral baixo (5%) e um nível de confiança alto (95%).

Após o processo de aplicação da amostragem, o número de unidades consumidoras foi reduzido de 13625 para 374, mantendo-se a mesma proporção de unidades consumidoras classificadas como regulares e irregulares por estrato (classe) do conjunto de dados original.

A Tabela 4 mostra o resultado da aplicação da amostragem estratificada das unidades consumidoras por estrato.

Tabela 4 - Amostragem estratificada das unidades consumidoras por estrato

| **CLASSE/FASE** | **REGULARES** | **IRREGULARES** | **TOTAL** |
|---|---|---|---|
| RESIDENCIAL | 320 | 33 | 353 |
| COMERCIAL | 16 | 1 | 17 |
| INDUSTRIAL | 1 | 0 | 1 |
| OUTROS | 3 | 0 | 3 |
| TOTAL | 340 | 34 | 374 |

Fonte: elaborado pelo autor

A Figura 12 ilustra os campos de um registro de consumo de energia elétrica de uma unidade consumidora na sua forma original, com um exemplo do conteúdo de cada campo.

Figura 12 - Campos de um registro de consumo de energia elétrica

| ID | DESC_UNIDADE_NEGOCIO | DESC_BAIRRO | COD_NUMERO_OPERATIVO | COD_ALIMENTADOR | IND_FASE_FORNECIMENTO |
|---|---|---|---|---|---|
| 23807 | PIRA - OESTE | CENTRO | 373867 | VOT07 | 2 |

| LAST_INSP1 | EV1 | COD_RUA | COD_CLASSE_CALCULO | IND_LOCALIZACAO | DESC_CLASSE_CALCULO | DT_REFERENCIA |
|---|---|---|---|---|---|---|
| 201309 | 1 | 2133601 | 1 | U | Residencial | 201309 |

| NUM_DIAS | QTD_ENERGIA_FATURADA | QTD_ENERGIA_REGISTRADA | CONSUMO_PAD | REG_IRREG |
|---|---|---|---|---|
| 32 | 206 | 206 | 193 | REGULAR |

Fonte: elaborado pelo autor

Os campos desses registros são descritos a seguir.

- *ID* – identificador da unidade consumidora
- *DESC_UNIDADE_NEGOCIO* – descrição da unidade de negócio
- *DESC_BAIRRO* – identificação do bairro da unidade consumidora
- *COD_NUMERO_OPERATIVO* – número do transformador
- *COD_ALIMENTADOR* – identifica o código do alimentador
- *IND_FASE_FORNECIMENTO* – fase de fornecimento (monofásico, bifásico ou trifásico)
- *LAST_INSP1* – data em que ocorreu a última inspeção na unidade consumidora
- *EV1* – Tipo de eletroválvula
- *COD_RUA* – identificador da rua da unidade consumidora
- *COD_CLASSE_CALCULO* – código da classe de consumo
- *IND_LOCALIZACAO* – urbano ou rural
- *DESC_CLASSE_CALCULO* – descrição da classe de cálculo
- DT*_REFERENCIA* – data de referência da medição do consumo

- *NUM_DIAS* – intervalo entre duas medições consecutivas
- *QTD_ENERGIA_FATURADA* – quantidade de energia faturada (kWh)
- *QTD_ENERGIA_REGISTRADA* – quantidade de energia registrada (kWh)
- *CONSUMO_PAD* – consumo de energia padronizado
- *REG_IRREG* – classificação da unidade consumidora após realização de inspeção por parte da concessionária.

## 5.4 Avaliação da efetividade do mecanismo proposto

Para a avaliação da efetividade do mecanismo, os dados de consumo de energia elétrica foram submetidos aos módulos de análise diferencial, global e combinação de Dempster-Shafer. Os detalhes mais importantes da implantação desses módulos estão descritos a seguir:

### 5.4.1 Análise diferencial

Na análise diferencial, para se determinar o perfil de comportamento local de uma unidade consumidora foi utilizado o atributo correspondente ao perfil de atividade, que se caracterizada pelo consumo mensal de energia elétrica de cada unidade (kWh).

Para verificar se a escolha do método estatístico exerceria alguma influência significativa no resultado da análise diferencial foram utilizados dois métodos estatísticos diferentes: média móvel simples e ponderada.

A média móvel é uma técnica utilizada para calcular a média dos n últimos valores de uma série cronológica (janela) e a média móvel é aplicada no cálculo do valor médio do consumo de energia elétrica, de cada unidade consumidora analisada na amostra.

A média móvel pode ser do tipo simples ou ponderado. O tipo simples representa a média aritmética dos valores de uma amostra n dos últimos períodos,

de modo que, à medida que ocorrem novas medições de consumo de energia elétrica, as medições mais antigas, que não se referem aos períodos considerados para o cálculo da média, são descartadas e novas médias móveis são calculadas considerando-se apenas os n períodos mais recentes.

No cálculo da média móvel simples, cada nova medição de consumo de energia elétrica de uma unidade incluída no cálculo tem exatamente o mesmo peso que as medições referentes aos períodos anteriores. Assim, quanto maior a quantidade de períodos utilizados para o cálculo da média, menor é a influência das medições mais recentes no resultado e mais lenta é a reação do indicador frente às novas medições de consumo de energia elétrica das unidades em questão.

Já no modelo de média móvel ponderada, as mais recentes medições de consumo de energia elétrica são ponderadas com peso maior do que as medições mais antigas. Ou seja, o cálculo da média móvel ponderada leva em consideração o peso atribuído a cada medição da amostra. Dessa forma, o mecanismo de detecção de fraudes pode reagir mais rapidamente a uma mudança do consumo.

Dentro dos modelos estatísticos, a média móvel é o modelo que tende a refletir de maneira mais adequada o perfil de comportamento de consumo de energia elétrica em função da sazonalidade que caracteriza a série analisada. Faz-se mister, no entanto, que o número de períodos incluídos no cálculo da média seja suficiente para que, teoricamente, o padrão volte a se repetir. Neste trabalho, o número de períodos na amostra analisada contempla doze meses – quantidade de ciclos suficientes para refletir a sazonalidade de consumo de energia de uma unidade consumidora.

A análise diferencial foi utilizada para determinar se o consumo de energia do mês de referência de uma determinada unidade consumidora diferia muito do valor de consumo médio, para os períodos considerados. O resultado da análise diferencial é uma variável denominada distância probabilística. A distância probabilística foi determinada por um modelo estatístico cujo resultado é um valor

que varia entre 0 e 0,5, quando a medição de consumo de energia elétrica do mês de referência for próximo à média de consumo da unidade consumidora e, entre 0,51 e 1, quando a medição for significativamente menor que a média do seu consumo.

Os valores da janela de observações e do fator e ponderação, arbitrados, empiricamente, após simulações com valores nas faixas de [1,15] e [0,1] e que deram os melhores resultados em termos de valor AUC foram 12 e 0.9, respectivamente.

### 5.4.2 Análise global

Como atributo global, foi escolhido o degrau de consumo em decorrência do fato de que, estatisticamente, as unidades consumidoras irregulares apresentam degrau de consumo. Além de se tratar de um atributo comumente utilizado pelas concessionárias de energia para selecionar unidades consumidoras suspeitas de fraude para inspeção. A escolha desse atributo deu-se com base em uma análise feita na base de dados, em que se constatou empiricamente que, de fato, boa parte das unidades consumidoras inspecionadas e identificadas a priori como sendo ou irregulares apresentavam degrau de consumo, ou seja, queda acentuada no consumo de energia elétrica, entre dois meses consecutivos.

A análise global foi utilizada para determinar se havia degrau de consumo acentuado entre dois meses consecutivos nas medições de consumo de energia das unidades analisadas. Assim como no módulo de análise local, o resultado da análise global é uma variável denominada distância probabilística. A distância probabilística foi determinada por um modelo estatístico cujo resultado é um valor que varia entre 0 e 0,5 quando o degrau de consumo entre dois meses consecutivos for próximo à média da unidade consumidora, e entre 0,51 e 1 quando o degrau de consumo for significativamente maior do que a média.

O resultado da análise global foi utilizado para fortalecer as evidências de fraude, determinadas pelo módulo de análise diferencial.

### 5.4.3 Combinação de Dempster-Shafer

Após a determinação dos resultados dos módulos de análise local e global, os resultados das análises diferencial e global foram então combinados utilizado a Teoria Matemática de Dempster-Shafer, gerando, assim, um escore final de evidencia de fraude (*m3*).

A Figura 13 ilustra um trecho de uma sequência de unidades consumidoras cujas medições de consumo de energia elétrica passaram pelas análises diferencial e global, juntamente com os resultados obtidos por meio da aplicação da Teoria Matemática de Dempster-Shafer (*m3*).

Na unidade consumidora identificada pelo ID 11922, verifica-se, por exemplo, que apesar da distância probabilística determinada pela análise diferencial ter indicado com 37,58% de probabilidade que a unidade consumidora apresentava irregularidades na medição do consumo de energia, a análise determinou com 100% de probabilidade que a unidade consumidora era suspeita de fraude em virtude da variação brusca de consumo de energia elétrica em 2 meses consecutivos.

Figura 13 - Trecho de uma sequência de medições de consumo de energia

| ID | IND_FASE_FORNECIMENTO | DESC_CLASSE_CALCULO | DT_REFERENCIA | CONSUMO_PAD | REG_IRREG | LOCAL | GLOBAL | *m3* |
|---|---|---|---|---|---|---|---|---|
| 465 | 1 | Residencial | 201408 | 235 | RREGULAR | 2,90% | 0,06% | 2,96% |
| 14428 | 2 | Residencial | 201408 | 309 | RREGULAR | 24,36% | 0,14% | 24,47% |
| 11922 | 2 | Residencial | 201408 | 461 | IRREGULAR | 37,58% | 100% | 100% |
| 1767 | 2 | Residencial | 201408 | 232 | RREGULAR | 23,46% | 0,15% | 23,58% |
| 1890 | 2 | Residencial | 201408 | 72 | RREGULAR | 33,10% | 0,13% | 33,18% |

Fonte: elaborado pelo autor

### 5.4.4 Avaliação de desempenho do mecanismo de detecção de fraudes proposto

Para avaliar o desempenho do mecanismo de detecção de fraudes proposto foram utilizadas as curvas ROC (*Receiver Operating Characteristics*) e AUC (*Area Under Curve*).

A curva ROC descreve o desempenho de um classificador utilizando duas dimensões (sensibilidade e especificidade). Ela permite verificar os comportamentos das taxas de verdadeiros positivos (TVP) e falsos positivos (TFP) em relação a vários níveis de limiar. Ela ajuda, ainda, na determinação do melhor ponto de operação de um detector, ou seja, do melhor valor do limiar que deverá ser utilizado para se obter a melhor relação entre a taxa de verdadeiro positivo (TVP) e a taxa de falso positivo (TFP) para a base de dados.

A curva AUC (*Area Under Curve*) foi utilizada para avaliar a capacidade de discriminação do mecanismo de detecção, ou seja, sua capacidade de classificar corretamente medições de consumo de energia elétrica analisadas.

Os valores de *m3* foram comparados com uma serie de limiares e, para cada valor de limiar, uma matriz de confusão foi gerada para determinar taxa de verdadeiros positivos e taxa de falsos positivos.

## 5.5. Síntese do capítulo

Este capítulo apresentou o método utilizado para validar a proposta. Neste capítulo discorreu-se sobre a escolha da técnica de amostragem mais adequada para que a massa de dados original fosse convertida em uma base de dados menor, sem, no entanto, perder a representatividade dos dados originais. No caso foi empregada a técnica de amostragem aleatória estratificada. Buscou-se detalhar os aspectos mais importantes da avaliação da efetividade do mecanismo proposto, com ênfase nos módulos de análise diferencial e global e combinação

de Dempster-Shafer. O desempenho do mecanismo de detecção de fraudes proposto foi avaliado utilizando-se as curvas ROC e AUC.

No próximo capítulo serão analisados os resultados decorrentes da validação da proposta.

# 6 ANÁLISE DOS RESULTADOS

Este capítulo apresenta uma análise crítica dos resultados obtidos nas simulações realizadas decorrentes da aplicação do mecanismo de detecção de fraudes proposto.

## 6.1 Considerações iniciais

Para avaliar o desempenho o mecanismo de detecção de fraude de consumo de energia elétrica proposto foram aplicados os dois modelos estatísticos considerados para a análise diferencial e a análise global, conforme descrito nos itens 5.4.1 e 5.4.2, utilizando-se uma amostra da base de dados de consumo de energia elétrica descrita no item 5.3.

O objetivo de utilizar dois métodos estatísticos foi verificar se a escolha do método estatístico, no conjunto de testes, teria algum efeito significativo nos resultados.

O resultado da combinação das duas análises (*m3*) foi comparado com os valores de limiares, variando de 1,0 até 0,0, com decrementos de 0,025.

Para cada valor de limiar foi gerada a matriz de confusão, por meio da qual calcularam-se os valores das taxas de verdadeiros positivos (TVP) e de falsos positivos (TFP) correspondentes a um determinado ponto da curva ROC. Após terem sido determinados os pontos da curva ROC, calculou-se o valor de AUC correspondente.

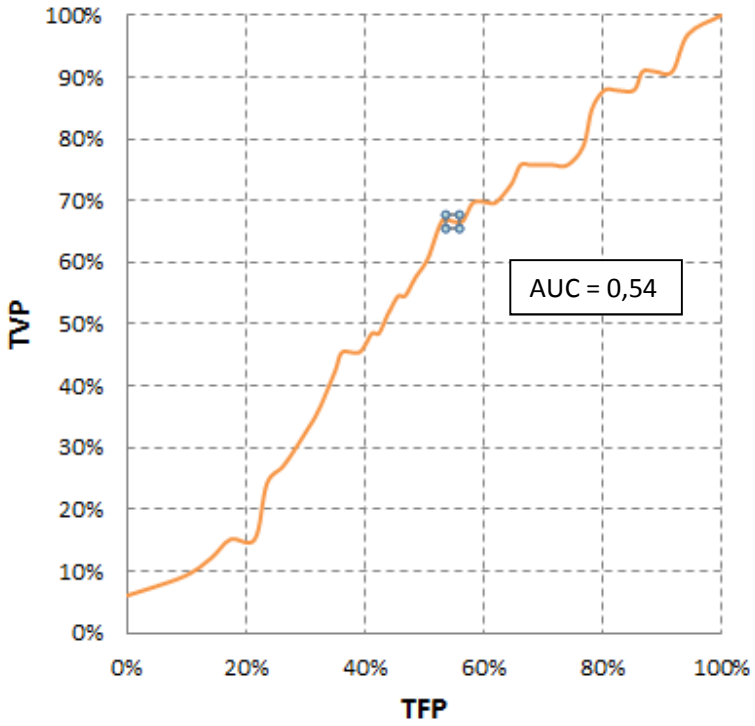
Para a avaliação de desempenho do detector proposto, os valores de *m3* foram comparados com uma serie de limiares e, para cada valor de limiar, uma matriz de confusão foi gerada para determinar taxa de verdadeiros positivos e taxa de falsos positivos.

## 6.2 Avaliação do desempenho do mecanismo proposto

A Figura 14 ilustra a curva ROC e o valor AUC resultante da aplicação do método estatístico baseado em média móvel simples sem a aplicação da análise global.

O melhor ponto de operação ocorre quando o valor do limiar é igual a 0,5250. Neste ponto têm-se TFP= 0,5294 e TVP = 0,6667.

Figura 14 - Curva ROC e AUC resultante da aplicação do método estatístico baseado em média móvel simples sem a aplicação da análise global

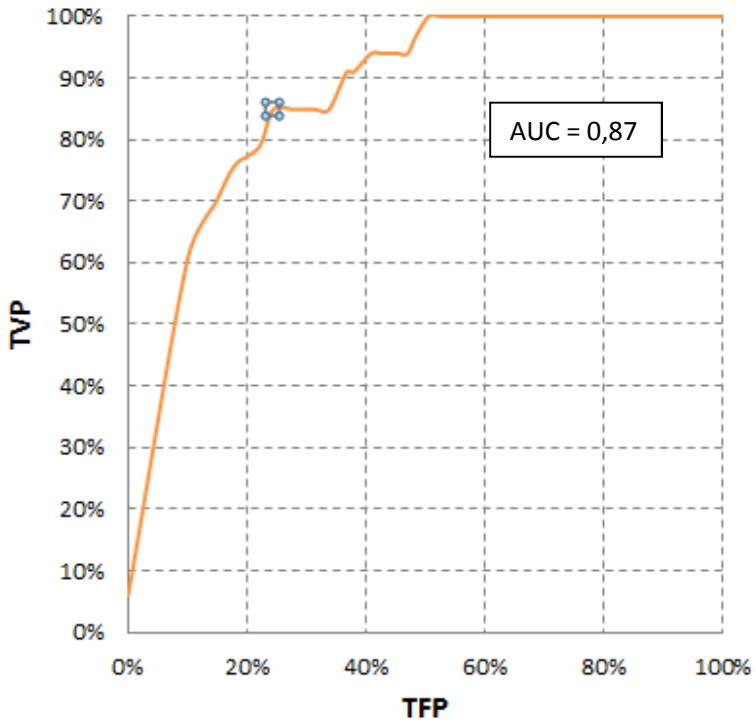



Fonte: elaborado pelo autor

A Figura 15 ilustra a curva ROC e o valor da AUC resultante da aplicação do método estatístico baseado em média móvel simples aplicando-se análise global.

O melhor ponto de operação ocorre quando o valor do limiar é igual a 0,875. Neste ponto têm-se TFP=0,2441 e TVP=0,8485.

Figura 15 - Curva ROC e AUC resultante da aplicação do método estatístico baseado em média móvel simples aplicando-se análise global



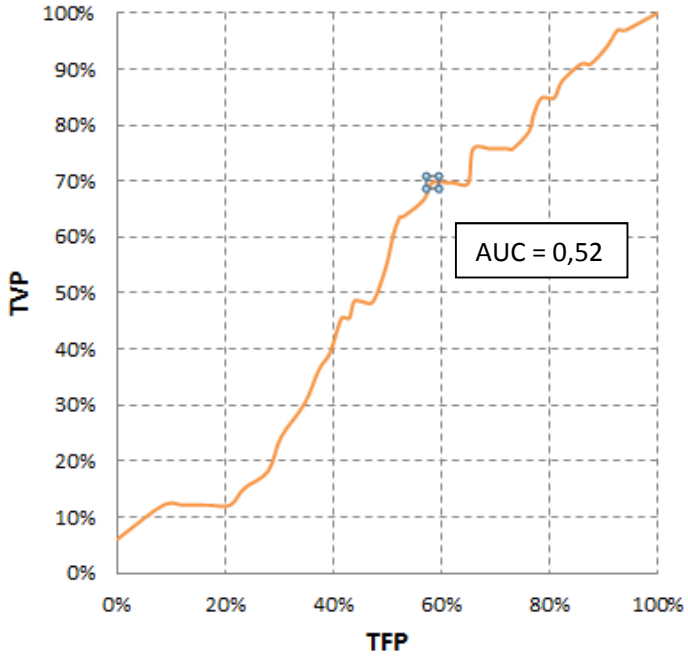

Fonte: elaborado pelo autor

Comparando-se as figuras 14 e 15, observa-se uma melhoria significativa da capacidade de discriminação do detector com a adição da análise global.

A Figura 16 ilustra a curva ROC e o valor de AUC resultante da aplicação do método estatístico baseado em média móvel ponderada na análise diferencial sem a aplicação da análise global.

O melhor ponto de operação ocorre quando o valor do limiar é igual a 0,475. Neste ponto têm-se TFP= 0,5824 e TVP = 0,6970.

Figura 16 - Curva ROC e AUC resultante da aplicação do método estatístico baseado em média móvel ponderada na análise diferencial sem a aplicação da análise global

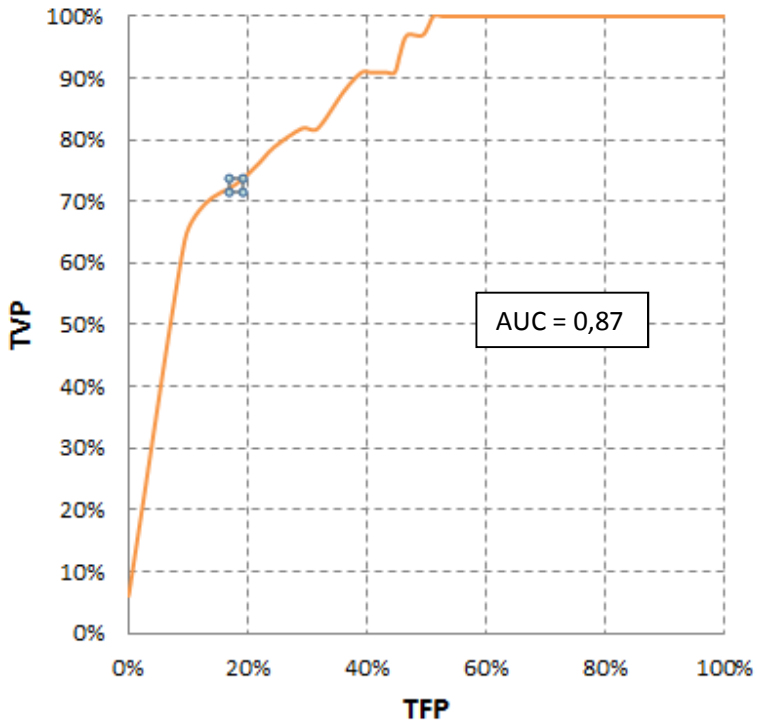



Fonte: elaborado pelo autor

A Figura 17 ilustra a curva ROC e o valor de AUC resultante da aplicação do método estatístico baseado em média móvel ponderada na análise diferencial, aplicando-se análise global.

O melhor ponto de operação ocorre quando o valor do limiar é igual a 0,925. Neste ponto têm-se TFP=0,1794 e TVP=0,7273.

Figura 17 - Curva ROC e AUC resultante da aplicação do método estatístico baseado em média móvel ponderada na análise diferencial, aplicando-se análise global




Fonte: elaborado pelo autor

## 6.3 Análise dos resultados

A Tabela 5 apresenta o resumo dos resultados obtidos em cada teste realizado.

Tabela 5 – Tabela comparativa dos resultados obtidos em cada teste

| Análise diferencial | Análise global | AUC | Limiar | TVP | TFP |
|---|---|---|---|---|---|
| Média móvel simples | Sem análise global | 0,54 | 0,5250 | 0,5294 | 0,6667 |
| | Com análise global | 0,87 | 0,8750 | 0,2441 | 0,8485 |
| Média móvel ponderada | Sem análise global | 0,52 | 0,4750 | 0,5824 | 0,6970 |
| | Com análise global | 0,87 | 0,9250 | 0,1794 | 0,7273 |

Fonte: elaborado pelo autor

Com base na análise dos resultados obtidos, constatou-se que os valores de AUC das curvas ROC permaneceram praticamente os mesmos, independentemente dos modelos estatísticos utilizados na análise diferencial, corroborando a ideia de que quando combinados com os resultados da análise global, os modelos estatísticos utilizados na análise local não exercem influência no resultado final.

Os dados mostram que, com a aplicação da análise global, os valores das curvas ROC saltaram com a utilização da média móvel simples e ponderada de 0,54 para 0,87 e 0,52 para 0,87, respectivamente, o que representa um ganho de pelo menos 61% na capacidade de discriminação do detector. Este resultado confirmou a hipótese inicial do trabalho de que a utilização de um atributo global melhora a capacidade de detecção.

## 6.4 Considerações acerca dos resultados

Os resultados obtidos, por meio dos testes realizados, mostraram que a utilização de atributos globais pode aumentar a capacidade de discriminação de um detector de fraudes. Por meio da utilização das curvas ROC e AUC verificou-se que independentemente do método estatístico utilizado para a abordagem local, a sua combinação com uma abordagem global melhora, significativamente, o desempenho do sistema de detecção de fraudes.

No entanto, algumas considerações devem ser feitas a respeito da abordagem utilizada na avaliação da efetividade do mecanismo proposto. A primeira diz respeito à utilização da média móvel para determinação do perfil de consumo de uma unidade consumidora com base no histórico de consumo.

Com base em uma análise empírica feita por amostragem na base de dados, constatou-se que, tratando-se de medição de consumo de energia elétrica, não existe um padrão regular de comportamento de uma unidade consumidora. Isso faz com que, muitas vezes, os perfis de consumos atípicos confundam-se com perfis de consumo típicos.

Portanto, não se pode afirmar que outro modelo estatístico, ainda que fornecesse um perfil de comportamento mais preciso da unidade consumidora, seria melhor. Isto ocorre porque, como a análise diferencial é baseada em desvio de comportamento, unidades consumidoras irregulares, que apresentassem históricos de consumo típicos de unidades consumidoras regulares, naturalmente passariam despercebidas de qualquer forma, independentemente do modelo estatístico utilizado. Esta observação evidencia a vantagem de adoção de outro método de detecção de fraude que, atuando em paralelo, poderia ser utilizado para reduzir a possibilidade de ocorrência de falso positivo.

A segunda consideração diz respeito à utilização do atributo global. O único atributo global utilizado neste trabalho foi o degrau de consumo. Essa escolha deveu-se ao fato de que, em geral, as unidades consumidoras que apresentam irregularidades no consumo de energia registram queda repentina no consumo de

energia elétrica. Isso não significa, entretanto, que unidades consumidoras regulares não possam registrar quedas atípicas. Um exemplo típico é o caso de variação de consumo de energia de unidades consumidoras da classe residencial em função de férias. Também podem ocorrer quedas acentuadas de consumo em função da diminuição da carga instalada junto à unidade consumidora ou mesmo em função de questões socioeconômicas, ambientais ou climáticas do local onde as unidades consumidoras encontram-se estabelecidas.

Por outro lado, também podem existir unidades consumidoras que, apesar de não apresentarem variações atípicas de consumo de energia e até mesmo de demonstrarem consumos de energia compatíveis ao de unidades consumidoras regulares, embora em números reduzidos, podem ser fontes de falsos positivos.

Como suporte para realização dos testes de efetividade do mecanismo de detecção foi utilizada uma amostragem aleatória estratificada. visando à diminuição da dimensionalidade e da discretização dos dados originais. A amostra gerada manteve a representatividade das unidades consumidoras quanto à sua classe de consumo e quantidade de fases, preservando a mesma proporção de unidades consideradas regulares e irregulares após a realização de inspeção pela concessionária.

No caso específico deste estudo, como buscou-se assegurar que o erro amostral fosse baixo e o nível de confiança alto. Dessa forma, pode-se afirmar que as conclusões obtidas a partir das análises efetuadas se aplicam ao conjunto de dados.

## 6.5. Síntese do capítulo

Este capítulo apresentou uma análise crítica acerca dos resultados obtidos na realização dos testes decorrentes da aplicação do mecanismo de detecção de fraudes proposto.

No próximo capítulo são apresentadas as principais conclusões acerca do trabalho e sugestões para trabalhos futuros.

# 7 CONCLUSÃO

O presente trabalho teve por objetivo propor um mecanismo de detecção de fraude de consumo de energia elétrica, utilizando dois métodos operando simultaneamente: um baseado em análise diferencial e o outro em análise global.

A análise diferencial foi utilizada para detectar desvios de perfil de consumo de energia elétrica das unidades com base na análise do consumo mensal (kWh). Para verificar se a escolha do método estatístico exerceria alguma influência significativa no resultado final da análise diferencial, foram utilizados dois métodos estatísticos (média móvel simples e média móvel ponderada).

Com base nos resultados observados, verificou-se que, dentro do cenário analisado, os dois métodos estatísticos utilizados para a análise diferencial apresentaram resultados próximos, com nenhuma influência no resultado final, quando combinados com os resultados da análise global.

O resultado da análise global, determinado com base no degrau de consumo, por sua vez, foi utilizado para fortalecer as evidências de fraudes determinadas pelo módulo de análise diferencial, baseado no consumo de energia mensal.

As evidências de suspeita de fraude determinadas por meio dessas duas abordagens foram então combinadas utilizando-se a Teoria Matemática de Dempster-Shafer para gerar um valor probabilístico de suspeita de fraude que, quando alcançado um limiar preestabelecido, dispara um alarme de alerta.

A validação da efetividade do mecanismo de detecção de fraudes de consumo de energia elétrica proposto a partir de uma amostra de dados de consumo de energia elétrica de unidades consumidoras já inspecionadas e identificadas, *a priori,* como sendo regulares ou irregulares, mostrou que o método foi apropriado para a finalidade pretendida.

A avaliação do desempenho do mecanismo de detecção de fraudes proposto foi feito por meio das curvas ROC e AUC e mostrou boa capacidade de discriminação e especificidade do detector.

Os resultados obtidos por meio da validação da proposta confirmaram a hipótese inicial de que, independente dos métodos estatísticos utilizados para determinação do perfil de consumo de uma unidade, a sua combinação com uma abordagem global baseada na análise do comportamento de unidades consumidoras irregulares melhora a capacidade de detecção de fraude. Mais precisamente, a utilização de uma abordagem global para fortalecer as evidências de suspeita de fraude determinada pela análise local propiciou um ganho mínimo de 61% na capacidade de discriminação do detector.

Por fim, ressalta-se que o mecanismo de detecção de fraudes proposto utilizou variáveis que apresentam aderência ao processo de detecção de irregularidade de consumo de energia elétrica, de tal forma que ele pode ser empregado, com os devidos ajustes ou alterações de parâmetros e variáveis, por quaisquer concessionárias de energia, esperando-se resultados compatíveis com os apresentados na seção 6.

## 7.1 Contribuições

O mecanismo proposto neste trabalho, além de contribuir com o equacionamento da detecção de fraudes de consumo de energia elétrica, serve como ponto de partida para trabalhos futuros na proposição de novos métodos de detecção de fraude, sob perspectivas mais amplas por meio da substituição ou integração de novos métodos de detecção de fraude no contexto das *Smart Grids*, ou mesmo em outros domínios de aplicação.

Além de se obter uma melhoria geral no que tange à detecção de fraude de consumo de energia elétrica, a possibilidade da combinação de diferentes evidências de fraudes apresentado pelo mecanismo proposto por este trabalho permitirá, que os outros modelos de detecção de fraudes já utilizados pelas concessionárias possam ser combinados entre si, propiciando maior diversificação de modelos e ampliando, dessa forma, a visão dos especialistas sobre o problema da detecção de fraude de consumo de energia elétrica.

## 7.2 Trabalhos futuros

Como os perfis dos consumidores são dinâmicos, seja pela mudança de comportamento natural do cliente, seja pela alteração da composição familiar (residenciais) ou econômica (comerciais e industriais); essa dinâmica demanda uma reavaliação periódica dos parâmetros adotados pelo sistema.

No caso específico do mecanismo proposto, a variável utilizada para determinação do perfil de comportamento da unidade consumidora foi o consumo de energia elétrica mensal (kWh). Entretanto, outras variáveis que apresentassem aderência ao processo de detecção de irregularidades na medição de consumo de energia elétrica poderiam ser determinadas, utilizando os mesmos modelos estatísticos (média móvel simples e ponderada) ou mesmo outros modelos, como z-escore. Além disso, poderia ser criado um atributo baseado na reputação da unidade consumidora.

No caso da análise global, uma sugestão seria: utilizar como atributos os dados adicionais compostos pelo Índice de Perda Comercial de um Alimentador ou por um Índice de Suspeita da Área que representasse a vulnerabilidade de uma determinada área às perdas comerciais, com base em indicadores socioeconômicos. Outra sugestão seria a utilização de variáveis baseadas em geoestatística.

Finalmente, como o mecanismo de detecção de fraude proposto neste trabalho permite a integração de novos métodos ou a substituição dos atuais por novos, trata-se de uma alternativa a ser explorada como forma de aumentar a capacidade de discriminação detecção de fraude. Assim sendo, outros modelos de detecção de fraude poderiam ser combinados.

Vale ressaltar, no entanto, que caso um novo módulo de detecção de fraude seja adicionado a um sistema existente, o ideal é que este novo módulo já forneça a evidência de fraude expressa em valores probabilísticos.

# REFERÊNCIAS

ABRADEE. **Redes de energia elétrica**. ABRADEE.2014. Disponível em: < http://www.abradee.com.br/setor-eletrico/redes-de-energia-eletrica > Acesso em: 20.fev.2014

ALBUQUERQUE, R. O. **Uma proposta de um modelo de confiança computacional para grupos em sistemas distribuídos**. Brasília, 2008. 171 p. Tese (Doutorado) – Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, 2008.

ANEEL. **Agencia Nacional de Energia Elétrica**. ANEEL.2014. Disponível em: <http://www.aneel.gov.br> Acesso em: 20.fev.2014.

AREAL, J. L. **Proposta de um modelo de confiança para o protocolo *Optimized Link State Routing* (OLSR)**. Brasília, 2008. 70 p. Dissertação (Mestrado) – Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, 2008.

AVEGLIANO, P. B. **Simulação da formação de parcerias entre agentes baseadas no conceito de reputação**. São Paulo, 2008. 85 p. Dissertação (Mestrado) – Departamento de Engenharia da Computação e Sistemas Digitais, Escola Politécnica da Universidade de São Paulo, São Paulo, 2008.

BENZI, V. M. **Proposta de um modelo de confiança para um sistema de gerenciamento de conteúdo de comércio eletrônico**. Brasília, 2011. 68 p. Dissertação (Mestrado) – Departamento de Engenharia Elétrica, Universidade de Brasília, Brasília, 2011.

BOGRAD, A. ***Trust and Reputation in the Smart Energy Grid***. Faculty of Mathematics and Natural Sciences of Rijksuniversiteit Groningen, 2012.

BROWN, R. E. ***Impact of Smart Grid on Distribution System Design***. IEEE Power Engineering Society General Meeting, p.1-4, 2008.

CHEN, Q.; AICKELIN, U., ***Anomaly detection using the Dempster-Shafer method***. In Proc. of the 2006 International Conference on Data Mining, DMIN 2006, pp.232-240, 2006.

CHIANG, W. J.; JOU, H.L.; WU, J. C. ***Maximum power point tracking method for the voltage-mode grid-connected inverter of photovoltaic generation system***. International Conference Sustainable Energy Technologies (ICSET), p.1-6, 2008.

CORTES, C.; PREGIBON, D., ***Signature-based methods for data streams***. Data Mining and Knowledge Discovery, p. 167-83, 2001.

CRUZ, C. C. P.; MOTTA, C. L. R. **Um modelo de sistema de reputação para comunidades virtuais**. In: SIMPÓSIO BRASILEIRO DE INFORMÁTICA NA EDUCAÇÃO, 17., 2006, Brasília, Brasil. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/sbie/2006/010.pdf>. Acesso em: 05 fev. 2013.

DASGUPTA, P. ***Trust as a Comodity***. In: Gambetta, Diego (ed.) Trust: Making and Breaking Cooperative Relations, Department of Sociology, University of Oxford, pp. 49-72, 2000.

DUARTE, O. C. M. B.; FERNANDES, N. C.; FERRAZ, L. H. G.; VELLOSO, P. B. **Um mecanismo de exclusão acurado baseado em confiança para controle de aceso em redes *ad hoc***. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 29., 2011, Campo Grande, Brasil. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/sbrc/2011/0028.pdf>. Acesso em: 05 fev. 2013.

DUARTE, O. C. M. B.; LAUFER, R. P.; PUJOLLE, G.; VELLOSO, P. B. **Análise de um modelo de confiança para redes móveis *ad hoc***. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 26., 2008, Rio de Janeiro, Brasil. Disponível em:

<http://www.lbd.dcc.ufmg.br:8080/colecoes/sbrc/2008/049.pdf>. Acesso em: 05 fev. 2013.

DUARTE, O. C. M. B.; LAUFER, R. P.; PUJOLLE, G.; VELLOSO, P. B. **Um novo modelo para confiança em redes *ad hoc***. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 24., 2006, Curitiba, Brasil. Disponível em: <http://www.lbd.dcc.ufmg.br:8080/colecoes/sbrc/2006/st17_4.pdf>. Acesso em: 05 fev. 2013.

EXAME. **O Brasil na onda das *smart grids*.** EXAME. Disponível em: http://exame.abril.com.br/revista-exame/edicoes/1040/noticias/o-brasil-na-onda-das-smart-grids> Acesso em 29.abril.2013

FALCÃO, D. M. ***Smart Grid* e Microredes: o futuro já é presente**. In: VIII Simpósio de automação de sistemas elétricos – SIMPASE, 2009, Rio de Janeiro, Brasil.

FARIA, L. T. **Sistema Inteligente Híbrido Intercomunicativo para Detecção de Perdas Comerciais**. Ilha Solteira, 2012. 112p. Dissertação (Mestrado) - Faculdade de Engenharia, UNESP, Ilha Solteira, 2012.

FAWCETT, T., ***An Introduction to ROC Analysis***, Pattern Recognition Letters, vol. 27, no. 8, pp. 861-874, 2006.

FAWECETT, T.; PROVOST, F., ***Adaptive Fraud Detection.*** Data Mining and Knowledge Discovery, Kluwe, 1, p. 291-316, 1997.

FERREIRA, P.; ALVES, R.; BELO, O.; CORTESAO, L., ***Establishing fraud detection patterns based on signatures***. Industrial conference on data mining, Leipzig, Germany, p.526-538, 2006.

FRAGA, J. S.; MELLO, E. R.; WANGHAM, M. S. **Uso de um modelo de confiança para a composição de serviços *web***. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 27., 2009,

Recife, Brasil. Disponível em: <http://www.lbd.dcc.ufmg.br/colecoes/sbrc/2009/060.pdf>. Acesso em: 05 fev. 2013.

GAMBETTA, D. ***Can We Trust Trust?*** In: Gambetta, Diego (ed.) Trust: Making and Breaking Cooperative Relations, Department of Sociology, University of Oxford, pp.213-237, 2000.

GUIMARÃES, P. H. V.;MURILLO, A.;ANDREONI, M.;MATTOS, D. M. F.;FERRAZ, L. H. G.;PINTO, F. A. V.;COSTA, L. H. M. K.;DUARTE, O. C. M. B**. Comunicação em Redes Elétricas Inteligentes: eficiência, confiabilidade, segurança e escalabilidade**. In: SIMPÓSIO BRASILEIRO DE REDES DE COMPUTADORES E SISTEMAS DISTRIBUÍDOS, 31., 2013, Brasília, Brasil. Disponível em: < http://sbrc2013.unb.br/files/anais/minicursos/minicurso-3.pdf >. Acesso em: 10 out.2013

HILAS, C. S.; SAHALOS, J. N., ***User profiling for fraud detection in telecommunications networks***. Proceedings of the 5th International Conference Technology and Autonation (ICTA'05), Thessaloniki Greece, p. 382-387, 2005.

JOSANG, A., ISMAIL, R. and BOYD, C. ***A Survey on Trust and Reputation Systems for Online Service Provision***. In: Decision Support Systems, USA, 2006.

KHURANA, H.; HADLEY, M.; NING, L.; FRINCKE, D.A. ***Smart-Grid Security Issues, Security & Privacy***. In: IEEE , v.8, n.1, p.81-85, 2010.

KOU, Y.; LU, C.; SIRWONGWATTANA, S.; HUANG, Y., ***Survey of Fraud Detection Techniques***, Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control, Taipei, Taiwan, March 21-23, 2004.

KOVACH, S. **Detecção de fraudes em transações financeiras via internet em tempo real**. São Paulo, 2011. 133 p. Tese (Doutorado) - Escola Politécnica da Universidade de São Paulo, São Paulo, 2011.

MAMANI, E. Z. S**. Cálculo de reputação em redes sociais a partir de dados da colaboração entre os participantes.** 2013. 73 p. Dissertação (Mestrado) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2013.

MANOEL, A. A. M. **Mecânica estatística de sistemas de reputação em redes autônomas**. São Paulo, 2012. 69 p. Dissertação (Mestrado) – Instituto de Física, Universidade de São Paulo, São Paulo, 2012.

MARSH, S. P. ***Formalizing Trust as a Computacional Concept***. Department of Computing Science and Mathematics, University of Stirling. Doctorate Thesis, 1994.

MATEI, I.; BARAS, J. S.; SRINIVASSAN, V., ***Trust-Based Multi-Agent Filtering for Increased Smart Grid Security*****.** Mediterranean Conference on Control & Automation (MED), Barcelona, Spain, 2012.

MORANDI, M.; ZULKERNINE, M., ***A Neutral Network Based System for Intrusion Detection and Classification of Attacks***, IEEE International Conference on Advances in Intelligent Systems - Theory and Applications, Luxembourg- Kirchberg, Luxembourg, November 15-18, 2004

NISTIR 7628. ***Guidelines for Smart Grid Cyber Security***. National Institute of Standards and Technology (NIST). v1.0, 2010.

nMentors. **Curso de Smart Grid**. nMentors. 2013. Disponível em: <http://www.nmentors.com.br/treinamentos/smart_grid.htm>Acesso em: 25 fev. 2013.

OLIVEIRA, R. D.; JÚNIOR, J. C. M. V. Benefícios e desafios de redes inteligentes. **Revista Eletrônica de Energia**, São Paulo, v. 2, n.1, p. 3-14, jan./dez. 2012.
Página Sustentável. **Medidores inteligentes são testados em Sete Lagoas**. Página Sustentável. 2013. Disponível em: <http://www.paginasustentavel.com.br/uploads/imagens/noticias/medidor_noticia.jpg> Acesso em: 15 out. 2013.

PATEL, J. ***A Trust and Reputation Model for Agent-Based Virtual Organization***. Thesis of Doctor of Philosophy. Faculty of Engineering and Applied Science. School of Eletronics and Computer Science. University of Southampton, 2007.

PINHEIRO JUNIOR, J. R. B. **Xenia**: um sistema de segurança para grades computacionais baseado em cadeias de confiança. 2008. 106 p. Tese (Doutorado) – Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, 2008.

PRADHAN, O.; AWAN, M.; NEWMAN, K.; BARNES, F. ***Trust and Reputation Approach to Smart Grid Security***. Department of Eletrical, Computer and Energy Engineering. University of Colorado, 2011

SILVA, J. D. S. **Uma abordagem híbrida por Dempster-Shafer e algoritmos genéricos para o problema de correspondência em estereoscopia**. Tese (Doutorado) - Ministério da Ciência e Tecnologia Instituto Nacional de Pesquisas Espaciais de São José dos Campos, São José dos Campos, 1999.

SILVA, V. B. **Um modelo de confiança certificado baseado em assinatura digital aplicado a sistemas multiagente**. Curitiba, 2009. 117 p. Dissertação (Mestrado) – Centro de Ciências Exatas e de Tecnologia, Pontifícia Universidade Católica do Paraná, Curitiba, 2009.

SINGH, R.; VATSA, M.; NOORE, A.; SINGH, S. K., ***Dempster Shafer Theory based Classifier Fusion for Improved Fingerprint Verification Performance***, Indian Conference on Computer Vision, Graphics and Image Processing, Springer, and Signal Processing, pp 1241-1244, may 1996.

SMART E-NERGY. **País diferente, diferentes desafios no *smart grid***. Grupo Editora Bolina, n.4 nov.dez. 2010. Disponível em: <http:// www.smartenergyonline.com.br>. Acesso em: 17 mar. 2013.

SMART GRID LIGHT. **Perguntas frequentes**. SMART GRID LIGHT. 2014. Disponível em: <http://smartgridlight.com.br/perguntas-frequentes> Acesso em: 20.fev.2014

TIPPER, D.; YAN, Y. Y.; QIAN, Y.; SHARIF, H. ***A Survey on Cyber Security for Smart Grid Communications*.** In: IEEE COMMUNICATIONS SURVEYS & TUTORIALS, v.14, n.4, p.998-1010, 2012.

TOLEDO, F. **Desvendando as redes elétricas inteligentes.** Rio de Janeiro: Brasport, 2012.

UCHOA, J. Q.; PANOTIM, S. M.; NICOLETTI. M. **Elementos da teoria da evidência de Dempster-Shafer**. Tutorial do Departamento de Computação da Universidade Federal de São Carlos. São Carlos, 2000.

ZUBEN, F. J. V. **Sistemas Baseados em Regras e Árvores de Decisão**. Notas de Aula. 2011. Disponível em: <ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ea072_2s11/topico6_EA072_2s11.pdf>. Acesso em: 20 fev. 2015.