XXIV SNPTEE
SEMINÁRIO NACIONAL DE PRODUÇÃO E
TRANSMISSÃO DE ENERGIA ELÉTRICA

22 a 25 de outubro de 2017
Curitiba - PR

CB/GTL/09

GRUPO - XV
GRUPO DE ESTUDO DE SISTEMA DE INFORMAÇÃO E TELECOMUNICAÇÃO PARA SISTEMAS ELÉTRICOS - GTL

SISTEMA COMPUTACIONAL PARA DETECÇÃO DE FRAUDES NO CONSUMO DE ENERGIA ELÉTRICA EM REDES DE DISTRIBUIÇÃO

CARVALHO, R. L. S (*) UFPA
VIEIRA, J. P. A UFPA
NUNES, M. V. A. UFPA
BEZERRA, U. H. UFPA
MONTEIRO, F. P. UFPA
OLIVEIRA, R. C. D CELPA

RESUMO

Este trabalho apresenta uma ferramenta computacional para detecção de fraudes no consumo de energia elétrica em unidades consumidoras (UCs) instaladas em redes de distribuição. A arquitetura do sistema computacional consiste de camadas formadas por módulos escritos em linguagens avançadas de programação (Python e Java), os quais são responsáveis por exercer funções específicas de adequação de base de dados, webservice, mineração de dados (técnicas de árvore de decisão), controle e gerenciamento de dados, e a interface com o usuário. A ferramenta é detalhadamente descrita e aplicada a bancos de dados reais de fiscalização e de consumo de energia elétrica.
.

PALAVRAS-CHAVE

Sistema Computacional, Mineração de dados, Perdas técnicas, Fraudes no consumo de energia, Redes de Distribuição

1.0 INTRODUÇÃO

Atualmente, o segmento de distribuição de energia elétrica é operado no Brasil por 67 concessionárias entre as quais nove estão localizadas na região norte do Brasil. Essas concessionárias enfrentam constantemente problemas relacionados ao aumento de perdas. Portanto, de acordo com Abradee(associação brasileira de distribuidores de energia elétrica) cerca de 63 distribuidoras de energia apresentam um percentual de 13,5% de perdas em relação à energia injetada no sistema global. As demais concessionárias apresentam índices acima de 20%. Grandes parcelas dessas perdas comerciais estão associadas principalmente devido ao furto, fraude, entre outros distúrbios na rede elétrica (1).

As perdas comerciais afetam tanto as concessionárias de energia, uma vez que diminuem a arrecadação, quanto às unidades consumidoras, devido à má qualidade da energia e os níveis precários de tensão. As perdas são consideradas com um dos principais fatores responsáveis por contribuir no aumento da tarifa de energia cobrada aos consumidores. A redução das perdas comerciais acarreta na maior segurança e rentabilidade, tanto para os consumidores quanto para as distribuidoras (1).

As perdas comerciais são altas por inúmeros fatores de cunho sócio-econômico e cultural tais como: desemprego, baixa renda, falta de habitação, infraestrutura insuficiente, preço elevado da energia e de acessórios de ligação e impunidade em relação à corrupção e à fraude, por exemplo.

(*) Rua Augusto Correa, n° 01 – CEAMAZON/UFPA Laboratório de Qualidade de Energia– CEP 66.075-100Belém, PA – BrasilTel: (+55 91) 981497468 – Email: rodcarvalhocomp@gmail.com

Os problemas relacionados às perdas comerciais são observados em diversos países, como: Malásia, Singapura, Taiwan, Espanha, América Latina, assim como no Brasil. Para a quantificação dessas perdas, as concessionárias idealmente deveriam realizar inspeções nos medidores de todas as Unidades Consumidoras (UC), quantificar todas as ligações clandestinas, além de executar o levantamento da iluminação pública de toda a área selecionada. No entanto, devido ao grande número de UC e ao alto custo das inspeções, a realização de tal prática torna-se inviável.

Por isso, torna-se importante a utilização de um sistema computacional para auxiliar a equipe de operação na detecção de fraude no consumo de energia elétrica em redes de distribuição. Com isso, o processo de seleção de consumidores a serem inspecionados será agilizado, aumentando a taxa de acertos de clientes com perdas comerciais e conseguintemente diminuindo os custos com inspeções.

Alguns trabalhos utilizaram técnicas de mineração de dados para a realização de análise e detecção de UC que apresentavam perdas comerciais. Dentre essas técnicas destaca-se o uso de ELM (*Extreme Learning Machine*), SVM (*Support Vector Machine*), Fuzzy, Algoritmos Evolucionários, Árvore de Decisão, entre outras. Para este trabalho, é importante ressaltar o uso da técnica de classificação Árvore de Decisão, por ser uma estrutura bastante utilizada na implementação de sistemas especialistas e também em problemas de classificação, devido a sua capacidade em aprender e paralelamente tomar decisões. Esse tipo de técnica toma como entrada uma situação descrita por um conjunto de atributos, que podem ser discretos ou contínuos, e retorna a uma decisão desejada.

No trabalho (2) são apresentados estudos para avaliação e detecção de fraudes no sistema elétrico, inicialmente, é realizado no sistema computacional uma pré-seleção de possíveis consumidores que apresentam fraudes no consumo de energia elétrica. Como resultados, apresenta-se uma taxa de acerto de classificação de fraudes igual a 40%. Já no trabalho (3) foi apresentado um estudo para a detecção de perdas não técnicas utilizando a mineração de dados com dados reais em uma indústria. Para isso, foram apresentados dois estimadores de estados para análise de tendências de consumo elétrico. Em (4) são descritos avanços nos métodos para minimizar as perdas não técnicas na mesma empresa, tendo como objetivo a detecção do perfil dos clientes com anomalia no consumo de energia, para isso, utilizou-se de redes Baysesian e árvore de decisão para a detecção dos tipos de padrões de perdas não técnicas.

Portanto, neste trabalho tem como objetivo apresentar uma ferramenta computacional capaz de identificar as unidades consumidoras (UC) que apresentam suspeitas de fraudes no consumo de energia elétrica. Para isso, serão utilizadas técnicas computacionais de mineração de dados.

A metodologia aplicada baseia-se na descoberta de conhecimento em base de dados (KDD -*Knowledge Discovery in Databases*), onde envolve uma intersecção de campos de pesquisa como aprendizagem de máquina, reconhecimento de padrões, bancos de dados, estatística, inteligência artificial, aquisição de conhecimento de sistemas especialistas, visualização de dados e computação de alto desempenho.

2.0 PERDAS DE ENERGIA ELÉTRICA

As perdas de energia em um sistema elétrico podem ser definidas como a diferença entre a energia fornecida a uma determinada rede elétrica e a energia entregue regularmente aos consumidores nessa mesma rede. Assim, as perdas globais nos sistemas de potência podem ser divididas em duas parcelas: perdas técnicas e perdas não-técnicas. As perdas técnicas ocorrem naturalmente nos sistemas elétricos, causadas por ações internas nos materiais, inerentes aos processos de transporte de energia e consistem principalmente na dissipação de energia nos diversos componentes dos sistemas elétricos, como condutores, transformadores, medidores e equipamentos elétricos (2). Essas perdas também podem ser estimadas por meio de modelos matemáticos como os existentes nos cálculos de fluxo de potência, onde é necessário conhecer a carga e as características do sistema elétrico. As perdas não-técnicas (ou mesmo perdas comerciais) são causadas normalmente por problemas relacionados a falta de faturamento energia distribuída, entre outros motivos. A seguir, são apresentados de maneira sucinta conceitos relacionados a essas perdas.

2.1 Perdas Técnicas

As perdas técnicas são causadas pelas próprias propriedades físicas dos componentes dos sistemas elétricos. Portanto, a mesma pode ser constituída como a parcela de energia que é perdida durante o processo de transporte de energia elétrica, sendo caracterizada pelo efeito Joule, que ocorre devido à própria resistência elétrica do condutor, sendo proporcional ao quadrado da corrente que passa pelo mesmo. Logo, todos os componentes que estão conectados aos sistemas elétricos apresentam perdas técnicas, podendo ser medidas ou estimadas por cálculo.

Dentre essas perdas, existem diversas outras perdas que normalmente são estimadas, como as perdas em conexões, devido o efeito corona, fuga nos isolamentos dos equipamentos e nas linhas de transmissão e

distribuição, perdas nos bancos de capacitores e reguladores de tensão, entre outros. Devido a dificuldade em calcular exatamente a contribuição dessas perdas, são estabelecidos valores estimados, normalmente entre 5% a 10% do valor total das perdas técnicas (2).

2.2 Perdas Não-Técnicas

As perdas não-técnicas, também conhecidas como perdas comerciais, são causadas principalmente devido aos problemas relacionados ao não faturamento da energia distribuída. Assim, apresentam-se como as principais causas comuns das perdas não-técnicas são: falta ou inexistência de medidores, falha ou falta de aferição dos medidores de energia, furto ou fraude da energia elétrica. Essas perdas podem ser compreendidas como a diferença entre as perdas globais e as perdas técnicas. Portanto, quaisquer perdas que não possam ser enquadradas como técnicas são comumente consideradas como perdas não-técnicas.
Dentre as maneiras que podem ser consideradas as perdas não-técnicas, neste trabalho destaca-se as perdas devido ao furto e a fraude de energia elétrica.

O furto de energia elétrica ocorre quando uma unidade consumidora é conectada diretamente a rede da distribuidora, sem anuência da concessionária ou mesmo permissão temporária. Portanto, são consideradas ligações irregulares, clandestinas, concentradas principalmente em áreas invadidas ou em favelas. A fraude de energia pode ser definida como o ato consciente de uma pessoa para eliminar ou reduzir a energia faturada, podendo ser considerada uma longa lista de possibilidades definida a partir dos casos de fraude encontrados. Ocorre também quando o medidor de energia é adulterado, ou quando é realizado um desvio no ramal de entrada entes do medidor. Nesses casos, o consumidor realiza um aumento de carga à revelia da concessionária em um circuito clandestino e na maioria das vezes de forma sofisticada (2).

3.0 MINERAÇÃO DE DADOS

Na metodologia para o desenvolvimento do algoritmo utilizou-se a biblioteca *sklearn* desenvolvida em *Python*. A linguagem de programação *Python*, estando amplamente estabelecida como uma das linguagens mais populares da comunidade científica, a mesma possui uma sintaxe clara e concisa, favorecendo a legibilidade do código fonte e conseguintemente tornando-a uma linguagem mais produtiva. Portanto, o *Python* é um *software* aberto, compatível com o GLP (*General PublicLicense*), permitindo que o mesmo seja incorporado a diversos produtos. Por ser uma linguagem de alto nível orientada a objeto e de tipagem dinâmica, forte e iterativa, torna-se uma escolha atraente para o desenvolvimento de algoritmos e análises exploratórias de dados (6)(7). A metodologia utilizada neste trabalho tem objetivo extrair conhecimentos e informações potencialmente úteis a partir de um grande volume de dados. Portanto, por meio dos processos do KDD será possível produzir conhecimento com intuito de identificar padrões, até mesmo os que eram considerados desconhecidos, visando um melhor entendimento a respeito dos problemas. A Figura 1 ilustra todos os processos pertencentes ao KDD.

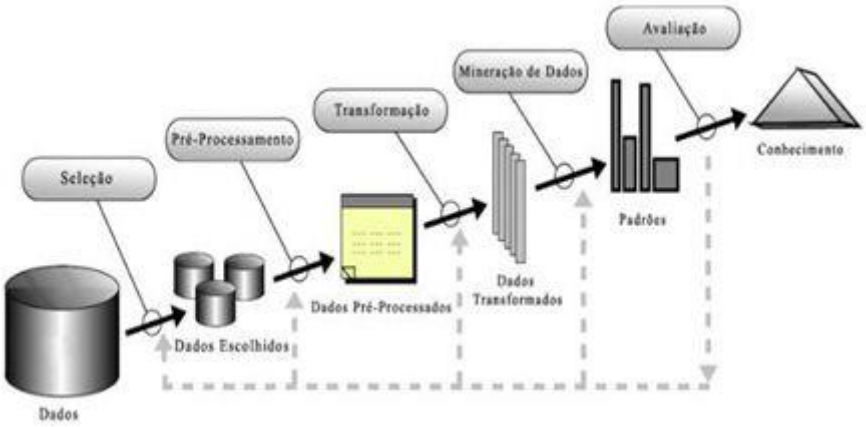

FIGURA 1 – Processo de KDD.

Dessa forma, o KDD envolve uma intersecção de campos de diversas pesquisas, como: Aprendizagem de Máquina, Reconhecimento de Padrões, Banco de Dados, Estatística, Inteligência Artificial, Aquisição de Conhecimento de Sistemas Especialistas, Visualização de Dados e Computação de Alto Desempenho. A metodologia utilizada no sistema computacional para a detecção de fraudes de energia em redes elétricas pode funcionar por meio da integração dos módulos de mineração e importação de dados.

4.0 SISTEMA COMPUTACIONAL

O sistema computacional foi desenvolvido utilizado o padrão de *software* MVC (Modelo Visão Controlador) dividindo as etapas do KDD entre os módulos. O *software* possui módulos específicos para as etapas de importação de dados, filtragem de dados, módulo de classificação e análise de resultados, por meio de gráficos com percentual das unidades consumidoras com suspeitas de fraudes e a localização geográfica das respectivas UC por meio do mapa da cidade.

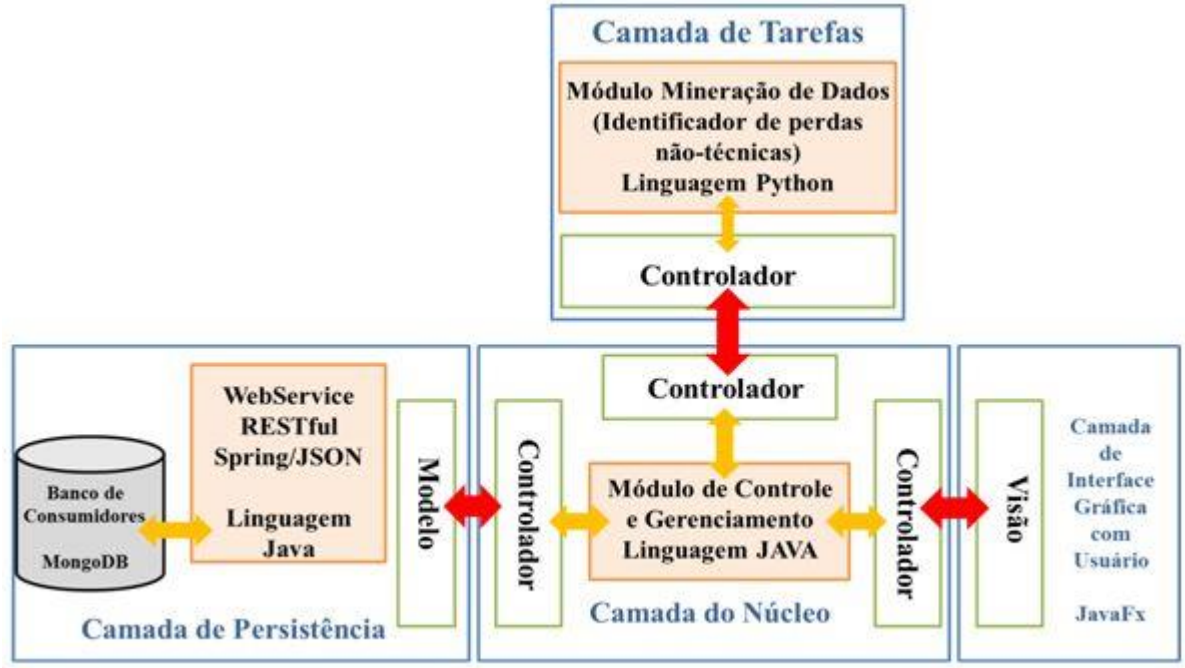

FIGURA 2– Fluxograma com a arquitetura do sistema computacional para detecção de fraudes.

Na Figura 2 é apresentado o fluxograma do sistema computacional utilizado para detecção de fraudes, sendo o mesmo dividido em quatro camadas: Camada de Tarefas, Camada de Persistência, Camada do Núcleo e a Camada de Interface Gráfica com o Usuário. Logo, cada uma destas tarefas é formada por módulos que são responsáveis por exercer determinadas funções. Com isso, é torna-se inevitável ressaltar o importante papel do MVC, permitindo que as camadas exerçam apenas as suas próprias funções de maneira independente. Com isso, não haverá problemas quanto à utilização ou fornecimento de informações, mesmo utilizando diferentes tipos de linguagem de programação entre os módulos, já que os mesmos podem até mesmo serem desenvolvidos avulso. Como principal responsável para que isso ocorra, é a Camada do Núcleo, pois a mesma tem o papel fundamenta de buscar cada informação solicitada pelo usuário e disponibilizar ao módulo responsável por exercer a tarefa, podendo fornecer os resultados, que já estão tratados e prontos a serem disponibilizados ao usuário.

4.1 Camada de Persistência

A Camada de Persistência, ou mesmo o CGD (Componente de Gerência de Dados) provê a infraestrutura básica para o armazenamento e a recuperação de objetos no sistema. Sua finalidade é isolar os impactos da tecnologia de gerenciamento de dados sobre a arquitetura do software (8). Nessa camada, encontra-se o módulo *Web Service RESTful*, onde o mesmo é utilizado como forma de integração e comunicação com o banco de dados e o sistema, de modo que o módulo de Controle e Gerenciamento possa realizar uma chamada para um serviço de qualquer banco de dados de forma transparente.

O *Web Service* desenvolvido foiimplementado em Java utilizando o *framework Spring* tento como formato de transição de informações escolhido JSON, pois é o formato usado pelo banco de dados *MongoDB*.Outro elemento da camada de persitencia é o banco de dados. Neste trabalho foi utilizado o *MongoDB* sequindo padrão *NoSQL*, sendo essa tecnologia utilizada para solucionar problemas gerados pelo âmbito do *big data*, promovendo variedades em soluções de armazenamento e processamento de grande volume de dados em aplicações, tendo a arquitetura de escalabilidade de forma horizontal, permitindo fornecer solução de inserções de novos dados de forma incremental e eficiente. O *MongoDB* é um banco de dados *schema-less*, orientado a documentos, sendo uma tecnologia de domínio público, conhecido como *open-source*, persistido os documentos no formato BSON (BinaryJSON), armazenando documentos nas chamadas coleções.Possui funções de agregração,exportação de dados nos formatos JSON e CSV,alta perfomance e alto desempenho para escritas.

4.2 Interface com Usuário

Normalmente, os *softwares* são desenvolvidos tendo como principal finalidade de serem utilizados por pessoas. Assim, apresentam como aspectos fundamental no projeto de sistemas é a interface com o usuário. O projeto da interface com usuário define uma forma de comunicação entre as pessoas e o sistema computacional. A interface define como um usuário comandará o sistema e como o sistema apresentará as informações ao usuário (9).

Com isso, a Figura 3 apresenta-se a interface do sistema computacional para detecção de fraudes em redes elétricas, tendo como finalidade facilitar a comunicação entre usuário e sistema computacional. Todas as opções para o usuário estão disponíveis na tela principal, tendo os recursos de localização georreferenciada, utilizando recursos de mapas para a melhoria de apresentação dos resultados e dividido em partes para especificar cada uma das particularidades dos módulos, com objetivo a facilitar a integração dos mesmos.

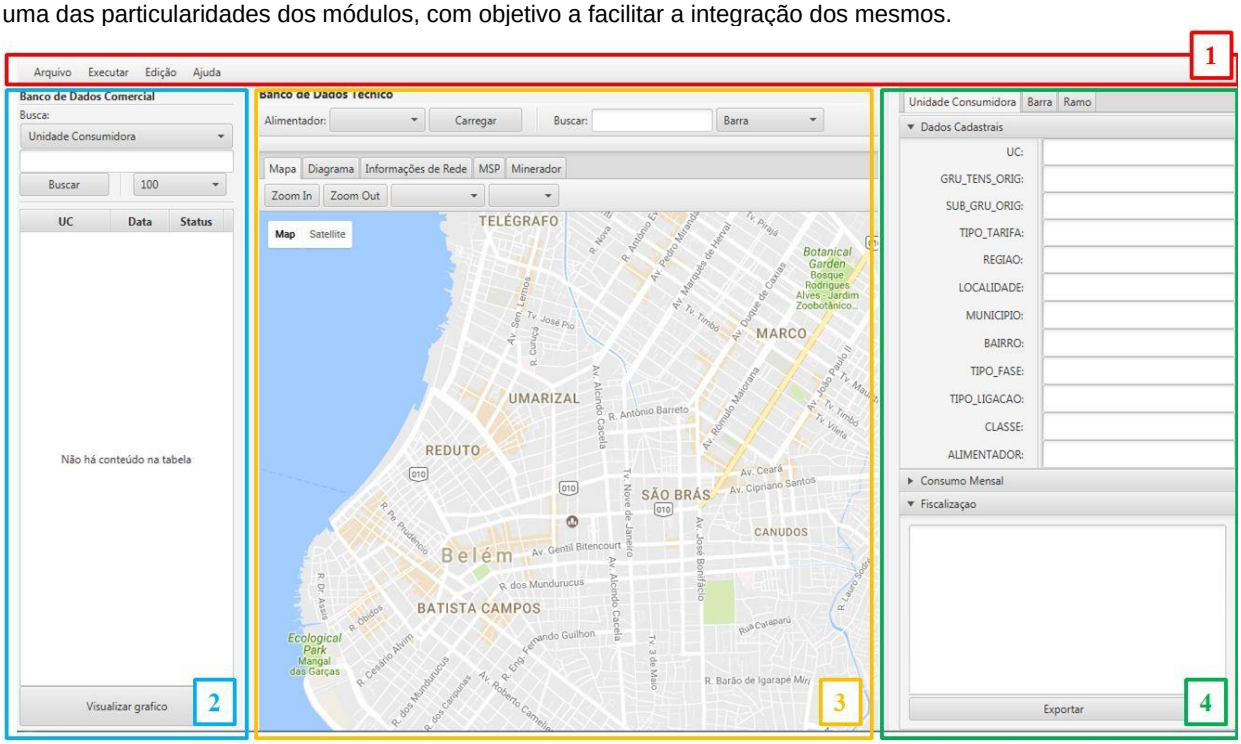

FIGURA 3– Interface do sistema computacional para detecção de fraudes.

Na Figura 3, a interface gráfica do sistema computacional para detecção de fraudes está dividida em 4 partes, ambas estando representadas por cores e números diferentes. Na parte superior da interface, representada pelo número pelo o número 1 e cor vermelha são disponibilizadas as tarefas, no menu “ARQUIVO” é possível acessar à interface gráfica do módulo de importação de dados, tanto para os dados oriundos da base de dados comercial e por fiscalização no local (*in loco*).

No menu “EXECUTAR” está disponível o módulo de detecção de fraudes do sistema computacional. Por último o menu “AJUDA” pode ser usado para fornecer algumas informações o usuário.

Na parte esquerda da interface, representada pela cor azul e o número 2, é possível gerenciar as informações das UCs pertencentes ao bairro ou alimentador específico, ou mesmo fazer a busca por UCs específicas. Assim, é possível apresentar todo o histórico de fiscalização de cada UC em forma de árvore do arquivo. Na parte central da interface, representada pela cor amarela e número 3 é apresentado a localização georreferenciada por meio do mapa da cidade, sendo ilustrado de forma iterativa e dinâmica para o usuário com a localização específica da UC disponibilizada pelo banco de dados. Na parte a direita da interface, representada pela cor verde e o número 4, o usuário visualizar as informações detalhadas e específicas das UC analisadas.

4.3 Mineração de Dados

A mineração de dados contida no sistema computacional, é composta basicamente por duas partes. A primeira, o módulo de importação está dividido em duas etapas: módulo de importação de dados dos consumidores e módulo de importação de dados da fiscalização. A segunda parte apresenta-se o módulo de mineração, contendo a ativação do módulo de mineração. Na Figura 4 está apresentado o fluxograma relativo ao processo do módulo de detecção de fraudes, também conhecido como Minerador de dados.

FIGURA 4–Módulo de detecção de fraudes.

O minerador de dados utilizado para a detecção de fraudes funciona por meio da integração do módulo de importação de dados e o módulo de mineração. No modulo de importação de dados, o fluxo está subdividido entre os dados dos consumidores que é importado da base do comercial da Empresa distribuidora do Estado do Pará, CELPA, onde nessa base contém informações das Unidades Consumidoras. A primeira etapa do processo de mineração é a formação das bases, que realiza o cruzamento das informações das UCs com o histórico de consumo com o histórico de fiscalização das UCs, onde serão armazenadas no banco de dados a partir das informações quer serão fornecidas pelo usuário, tendo a tarefa de alimentar a base a partir dos módulos de importação específica.

Os dados da base comercial contêm informações do consumo mensal das UCs e informações do tipo de ligação, grupo de faturamento, grupo de tensão, classe de consumo da UC, bairro, algumas UCs com coordenadas georreferenciadas. Após a importação dessas informações e criado no banco de dados nas medias estatísticas com o desvio padrão para cada mês e a média de consumo por mês para cada UC.

Na importação de dados da base de dados de fiscalização da empresa CELPA, os dados contêm a seleção da base de dados e a parte de validação de atributos apresenta-se como padrões os seguintes:

a. Código da Unidade Consumidora: Número específico da conta ou contrato;
b. Data do serviço de fiscalização: Data da inspeção realizada no local (*in loco*);
c. Código do serviço: Código do serviço realizado;
d. O *label* do estado da UC: apresentam-se informações, 0: para o caso em que a UC esteja normal e 1: para o caso em que a UC apresenta algum tipo de irregularidade;
e. Troca do medidor: terá as informações 0: para o caso em que o medidor não tenha sido trocado e 1: para o caso que o medidor já tenha sido substituído.

A etapa do módulo do minerador poderá ser utilizada pelo usuário, permitindo realizar a classificação da unidade consumidora e fornecendo as informações contidas no banco de dados, como: a subdivisão por bairro, subgrupo de tensão de origem, grupo de tensão de faturamento, tipo de fase, tipo de ligação e classe de consumo. A partir de toda essa classificação, a base de dados para o processo de mineração pode ser criada, podendo ser especificada de acordo com o reconhecimento e características desejadas da região a ser analisada. A Figura 5 apresenta a tela do módulo de mineração de dados do sistema computacional.

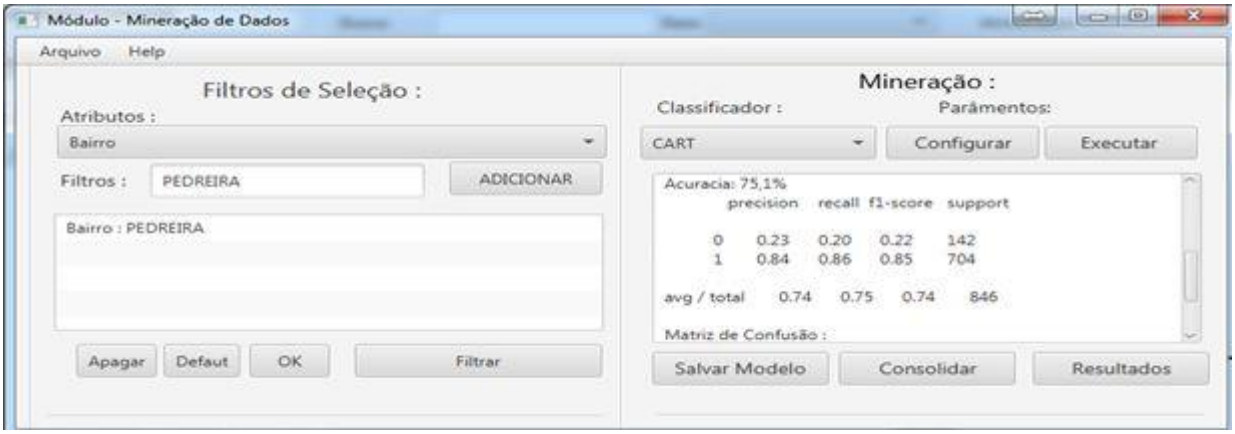

FIGURA 5– Tela do Módulo de Mineração de Dados.

Após a realização da seleção dos filtros, que posteriormente serão utilizados para a criação da base de dados, o processo do classificador é habilitado, com isso, o usuário poderá escolher uma das inteligências de classificação a serem utilizadas no processo do minerador de dados, sendo estas as árvores de decisões *Randon Forest* ou *Cart*. Essa alteração dos parâmetros para escolha pode ser realizada por meio da opção disponível na tela "CONFIGURAR".

Posteriormente, será disponibilizada ao usuário a taxa de acurácia do classificador, que estabelece a precisão e exatidão dos dados e informações fornecidas pelo minerador de dados. Também será fornecida a matriz de confusão, que normalmente é utilizada como ponto de partida para uma série de técnicas estatísticas, sendo esta, uma matriz quadrada que expressa o número das unidades de amostra distribuído a uma categoria particular relativa uma categoria atual (SUAREZ, 2012).

Após a primeira análise do classificador, caso o usuário considere o resultado da classificação satisfatório, o mesmo poderá salvar este modelo como preferencial e assim poder utilizar novamente em outros filtros de seleção. Assim, poderá consolidar a classificação que ficará salva no banco de dados, podendo visualizar os resultados por meio de tabelas e de localizações geográfica por meio de mapas.

5.0 RESULTADOS

O sistema possui três formas de apresentação de resultados para auxiliar usuário nas escolhas das UCs que serão fiscalizadas, para demonstração dessas opções que estão disponíveis foi executado análise alimentador PD-05 do bairro da Pedreira tento esse alimentador georreferenciamento.

Após o processo de seleção dos filtros e execução do minerador apresentado na seção anterior, o usuário poderá seguir para análise da classificação da UCs.

Na classificação das UCs do PD-05, usando árvore de decisão *Cart,* com seguintes paramentos, Tabela 01, a taxa de acurácia foi de 75,1% :

Tabela 01: Paramentos da Árvore de Decisão

criterion='gini'	splitter='best'	max_depth=None	min_samples_split=2	minsamples_leaf=1
max_features=None	random_state=None	max_leaf_nodes=None	class_weight=None	presort=False

Na Figura 6, mostra a lista das UCs rotuladas após o processo de classificação do minerador com o label de "Fraude" ou como "Regular". Sendo possível o usuário exportar o resultado da lista.

Outra opção de analise disponível para o usuário mostra a distribuição da classificação de fraude por bairro após a classificação. Outra forma de resultados está contida na aba "gráfico". Aonde o usuário pode verificar a quantidade das UCs que foram classificadas como fraudadora.

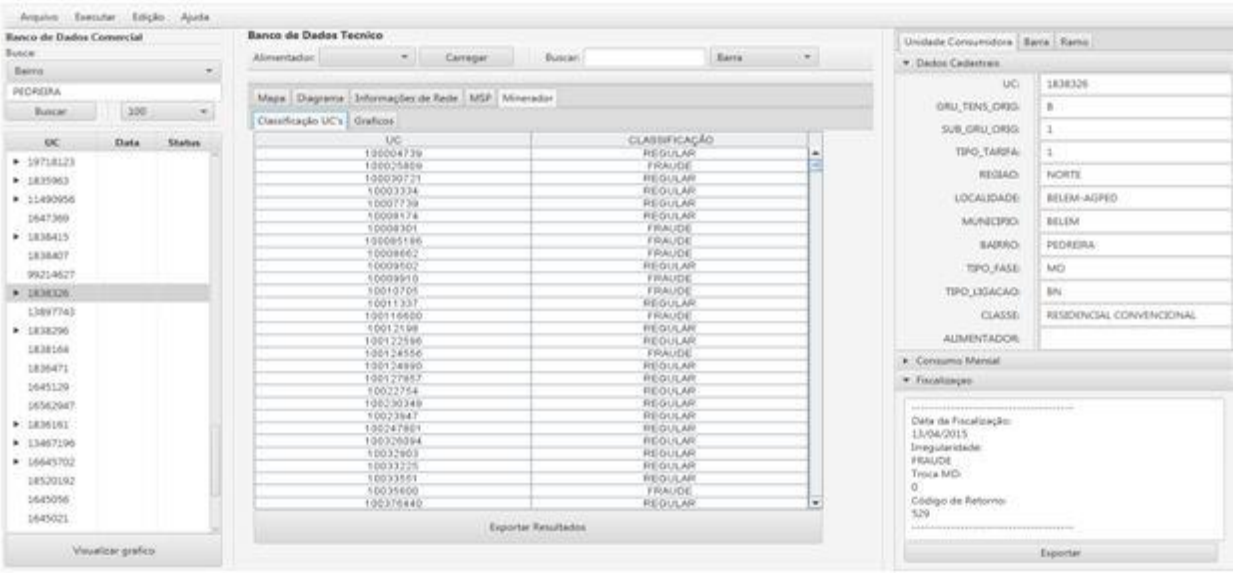

Figura 6. Tela de Resultados da Classificação.

Outra forma de gestão dos resultados mais iterativa para usuário é a demonstração das UCs através do mapa da cidade. O sistema disponibiliza para usuário opção de visualização de localização da UCs na Figura 7.

Figura 7. Tela com Unidades Consumidora Fiscalizadas

Na Figura 7, do lado esquerdo, o usuário pode escolher entre três formas de demonstrar as UCs, sendo através dos bairros, o alimentador especifico ou uma UC especifica. Caso o usuário "clique" em uma das UCs listadas no lado direito apresenta os dados gerais da UCs e também o consumo mensal dos últimos anos.
No mapa central são mostradas as UC que cometeram fraude comprovadas com fiscalização *in loco*, em cor vermelha, em azul as UCs classificadas como regular pela inspeção *in loco* e em verde as UCs classificadas como fraudadora pelo minerador de dados.

6.0 CONCLUSÃO

Nesse trabalho foi apresentado um sistema computacional desenvolvido para detecção de fraudes em redes de distribuição da Celpa, com objetivo de reduzir as perdas comerciais da distribuidora.

Um fator que ajudar a redução das perdas nas empresas de distribuição é combate *in loco* das UCs fraudadoras, otimização do processo de escolhas das UCs fiscalizadas unido com conhecimento do especialista torna o processo de fiscalização mais eficiente, para isso o sistema apresenta interface com várias opções de gerenciamentos de resultados.

Nesse trabalho foi desenvolvido toda uma camada de gerenciamento de dados, através de criação do banco de dados com Mongodb utilizando tecnologia NoSQL, com características de big data, tento alto desempenho nas funções para agregação de informação, os processos estatísticos estão automatizados através desse banco de dados. Onde fica condicionado ao usuário informações de todas as fiscalizações realizadas *in loco*, como o histórico dos consumos das UCs.

No processo de classificação das UCs foram implementados dois algoritmos de árvore de decisão tendo as taxas de valor de acurácia acima de 70%, possuindo taxa de acerto considerado alto na literatura. É possível futuramente extrair conhecimentos técnicos com estudos sobre as regras das técnicas de árvore, que estão disponíveis para usuário.

O sistema apresenta interface com diversos recursos, localização das UCs em mapas, lista para fiscalização, gráficos para análise por bairro, todos esses processos são otimizados e combate as perdas técnicas na rede da distribuidora.

7.0 REFERÊNCIAS BIBLIOGRÁFICAS

(1) http://www.abradee.com.br/setor-de-distribuicao/perdas/furto-e-fraude-de-energia. acesso <03/04/2017>.

(2)J. Filho, “Fraud identification in electricity company costumers using decision tree,” in Proc. IEEE/PES Int. Conf. Systems, Man and Cybernetics,The Hague, The Netherlands, 2004.

(3)C.León, F.Biscarri, I.Monedero, J.I.Guerrero, J.Biscarri, and R.Millán, “Variability and trend-base generalized rule induction model to NTL detection in power companies,”IEEE Trans. Power Syst., vol.26, no. 4, pp. 1798–1807, Nov. 2011.

(4)I. Monedero, F. Biscarri, C. León, J. I. Guerrero, J. Biscarri, and R. Millán, “Detection of frauds and other non-technical losses in a power utility using Pearson coefficient, Bayesian networks and decision trees”, Electrical Power and Energy Systems, Vol. 34, p-90-98, 2012.

(5) PENIN, C. A. S. “Combate, Prevenção e Otimização das Perdas Comerciais de Energia Elétrica”. 2008. (Tese de Doutorado). Escola Politécnica da Universidade de São Paulo. São Paulo – SP, Brasil, 2008.

(6)P.F. Dubois, editor. *Python: Batteries Included*, volume 9 of *Computing in Science &Engineering*.IEEE/AIP, May 2007.
(7)K.J. Milmann and M. Avaizis, editors. *Scientific Python*, volume 11 of *Computing inScience & Engineering*. IEEE/AIP, March 2011.

(8)COAD, P.; Yourdon, E.. Projeto Baseado em Objetos. Editora Campus: 1993.

(9)FALBO R, A. Projeto de Sistemas de Software. “Notas de Aula”. 2011.

(10)SUAREZ, A. F.; CANDEIAS, A. L. B. “Avaliação de Acurácia da Classificação de Dados de Sensoriamento Remoto para o município de Maragogipe”. IV Simpósio Brasileiro de Ciências Geodésicas e Tecnologias da Geoinformação Recife - PE, 2012

8.0 DADOS BIOGRÁFICOS

Rodrigo Luis da Silva Carvalho Graduado pelo Instituto de Estudos Superiores da Amazônia em Engenharia de Computação no ano de 2015. Atualmente mestrando em Sistema de Energia pelo Programa de Pós-Graduação em Engenharia Elétrica (PPGEE) da Universidade Federal do Pará. Atua em P&D no Centro de Excelência em Eficiência Energética da Amazônia (CEAMAZON).

João Paulo Abreu Vieira Possui graduação (2004), mestrado (2006) e doutorado (2009) em Engenharia Elétrica pela Universidade Federal do Pará. Atualmente é professor adjunto da Faculdade de Engenharia Elétrica da Universidade Federal do Pará. Seus interesses em pesquisa estão focados em modelagem, simulação, estabilidade, controle e segurança de sistemas elétricos de potência; integração da geração distribuída em sistemas de energia elétrica. É membro do IEEE PES.

Flavia Pessoa Monteiro Ingressou no doutorado em Engenharia Elétrica na área de Computação Aplicada pela Universidade Federal do Pará no Programa de Pós-Graduação em Engenharia Elétrica - PPGEE, no primeiro semestre de 2015. Bacharel em Engenharia da Computação pelo Instituto de Ensino Superior da Amazônia(IESAM) no ano de 2014. Mestre em Engenharia Elétrica com área de concentração em Computação Aplicada com linha de pesquisa em Inteligência Computacional. Atualmente possui linha de pesquisa na área de tecnológias sócias aplicando conhecimentos de Inteligência Computacional e Big Data.

Marcus Vinicius Alves Nunes Possui Mestrado em Engenharia Elétrica pela Universidade Federal do Pará (UFPA) em 1996 e Doutorado em Engenharia Elétrica pela Universidade Federal de Santa Catarina (2003). Atualmente é Professor Associado da Universidade Federal do Pará, e Coordenador do Laboratório de Alta e Extra Alta Tensão da UFPA. Vice-Chair do IEEE (Institute of Electrical and Electronics Engineers) Seção Centro Norte do Brasil. Possui experiência na área de Sistemas Elétricos de Potência, com ênfase em Máquinas Elétricas, conversores estáticos e Dispositivos de Potência, Fontes Renováveis de energia, Dispositivos de Alta Tensão, atuando principalmente nas seguintes linhas de pesquisa: integração de aerogeradores e sistemas fotovoltaicos em redes elétricas, geração distribuída, geradores de indução duplamente excitados, geradores síncronos direct drive, controle inteligente aplicado a sistemas elétricos de potência.

Ubiratan Holanda Bezerra Possui graduação em Engenharia Elétrica Opção Eletrônica pela Universidade Federal do Pará (1976), mestrado em Engenharia Elétrica pela Universidade Federal de Itajubá (1980) e doutorado em Engenharia Elétrica pela Universidade Federal do Rio de Janeiro (1985). Atualmente é professor titular da Universidade Federal do Pará. Tem experiência na área de Engenharia Elétrica, com ênfase em Geração da Energia Elétrica, atuando principalmente nos seguintes temas: sistemas de potência, qualidade de energia, wind generation, power quality e geração eólica.

Rafael Cavalcante de Oliveira Graduado pela Universidade Federal do Pará em Engenharia Elétrica no ano de 2009. Atualmente Engenheiro de Operações nas Centrais Elétricas do Pará, gerente do projeto de pesquisa em conjunto com a UFPA e cursando o mestrado em sistemas de energia pelo PPGEE da UFPA