

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
FAENG – FACULDADES DE ENGENHARIA, ARQUITETURA E
URBANISMO E GEOGRAFIA
ENGENHARIA ELÉTRICA

MARIANA VIUDES VILLALBA

METODOLOGIA EMPREGADA PELA CONCESSIONÁRIA DE ENERGIA
ELÉTRICA ENERSUL PARA DETECÇÃO DE FRAUDES NA MEDIÇÃO
DE ENERGIA ELÉTRICA

CAMPO GRANDE
2014

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL
FAENG – FACULDADES DE ENGENHARIA, ARQUITETURA E
URBANISMO E GEOGRAFIA
ENGENHARIA ELÉTRICA

MARIANA VIUDES VILLALBA

METODOLOGIA EMPREGADA PELA CONCESSIONÁRIA DE ENERGIA
ELÉTRICA ENERSUL PARA DETECÇÃO DE FRAUDES NA MEDIÇÃO
DE ENERGIA ELÉTRICA

Trabalho apresentado à Universidade Federal de Mato Grosso do Sul,
Curso de Graduação em Engenharia Elétrica, como requisito parcial
para aprovação na disciplina Trabalho de Conclusão de Curso II.
Orientador: Prof. Dr. Luigi Galotto Junior

CAMPO GRANDE
2014

MARIANA VIUDES VILLALBA

METODOLOGIA EMPREGADA PELA CONCESSIONÁRIA DE ENERGIA
ELÉTRICA ENERSUL PARA DETECÇÃO DE FRAUDES NA MEDIÇÃO
DE ENERGIA ELÉTRICA

Este trabalho foi julgado adequado e aprovado para a obtenção do título de graduação em
Engenharia Elétrica da Universidade Federal de Mato Grosso do Sul

Campo Grande, 19 de dezembro de 2014

Prof. Dr. Jéferson Meneguín Ortega
Coordenador do Curso de Engenharia Elétrica

BANCA EXAMINADORA:

Prof. Dr. Luigi Galotto Junior
Universidade Federal de Mato Grosso do Sul
Orientador

Prof. Dr. Evandro Mazina Martins
Universidade Federal de Mato Grosso do Sul
Banca

Eng. Eletricista João Ricardo Costa do Nascimento
Universidade Federal de Mato Grosso do Sul
Banca

UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL

DECLARAÇÃO DE AUTORIA E AUTENTICIDADE

Eu Mariana Viudes Villalba residente e domiciliada na cidade de Campo Grande, Estado do Mato Grosso do Sul, portadora do RG de nº 0938240041 – MEX/MS e CPF nº 030.961.771-55, declaro que o Trabalho de Conclusão de Curso (TCC) apresentado, com o título “Metodologia empregada pela concessionária de energia elétrica Enersul para detecção de fraudes no sistema de medição de energia elétrica” é de minha autoria e assumo a total responsabilidade pelo seu conteúdo e pela originalidade do texto. Declaro que identifiquei e referenciei todas as fontes e informações gerais que foram utilizadas para construção do presente texto. Declaro também que este artigo não foi publicado, em parte, na íntegra ou conteúdo similar em outros meios de comunicação, tendo sido enviado com exclusividade para a conclusão do curso de graduação em Engenharia Elétrica da Universidade Federal de Mato Grosso do Sul (UFMS).

Campo Grande, 19 de dezembro de 2014.

Mariana Viudes Villalba

RA 2008.210.30-330

Curso de Engenharia Elétrica – UFMS

**Dedico aos meus pais pelo amor,
cuidado e paciência.**

AGRADECIMENTOS

Agradeço a Deus pelo cuidado e por sempre me manter de pé, iluminando meu caminho e me concedendo forças para continuar.

Aos meus pais que me apoiaram durante todos esses anos, com muita paciência e muito cuidado, sempre me incentivando e acreditando em mim.

Ao meu professor orientador Dr. Luigi Galotto Junior, que prontamente se disponibilizou a me ajudar e se manteve presente em todos os momentos que precisei.

À equipe Enersul, que prontamente contribui para a realização deste trabalho fornecendo informações e me dando liberdade para trabalhar. Obrigada ao meu gerente Paulo Roberto, João Ricardo, Tiago Lucas, Johan, Silmeire, Paula, Bruna e a todos os colaboradores do Departamento de Medição e Combate às Perdas Comerciais.

E finalmente, aos meus amigos que de tão chegados transformaram-se em irmãos, sempre me apoiando e fazendo meus dias mais felizes. Obrigada Rogério, Juliana, Ana Cristina, Wellington, Ricardo e Pedro. Amo vocês!

**“Um povo que não lê a Bíblia é,
sem dúvida, propenso à fraude, ao
infortúnio e ao sofrimento.”**

Hermes Machado

RESUMO

As perdas comerciais de energia elétrica tem sido a grande preocupação das concessionárias, sendo cada vez mais difícil alcançar a meta regulatória.

A concessionária de energia, Enersul, trabalha com inúmeras ferramentas tecnológicas, que auxiliam na detecção de irregularidades na medição. A proposta da utilização de novas tecnologias é a diminuição do número de fiscalizações e consequente diminuição dos gastos com investimentos em perdas comerciais.

O algoritmo do DNA da fraude é uma ferramenta que trabalha com várias variáveis decodificadas, todas compactadas em cinco camadas distintas. Estas camadas são somadas e o resultado é denominado score. Unidades consumidoras com valores de score acima de 200 são consideradas suspeitas, no entanto é preciso fazer uma análise mais detalhada do perfil de consumo e das características da UC.

Inúmeras variáveis compõem o algoritmo DNA da fraude, desde informações trazidas pelo leiturista, até medições feitas nos alimentadores e transformadores de distribuição. Cada variável é denominada preditora. Cada uma delas têm um peso associado, que retrata o nível de importância daquela variável na detecção da irregularidade.

Após todas as preditoras serem classificadas com seus respectivos pesos, elas são agrupadas em cinco camadas distintas e cada camada recebe um peso correspondente à sua importância no resultado final. Como não existem registros na concessionária de como estes pesos foram encontrados, pode-se concluir que inúmeros processos podem ter sido utilizados na determinação dos mesmos. Este trabalho retrata uma das alternativas que podem ter sido usadas na obtenção dos valores dos pesos de cada camada. A hipótese considerada foi o uso de regressão linear simples para determinação dos melhores valores atribuídos.

Palavras-chave: Perdas comerciais, DNA da fraude, score, decodificadas, compactadas, regressão linear.

ABSTRACT

The commercial losses of electricity has been the major concern of dealers, being increasingly difficult to achieve regulatory goal.

The power utility, Enersul, working with numerous technological tools that assist in the detection of irregularities in the measurement. The proposed use of new technology is reducing the number of inspections and consequent reduction of spending on investments in commercial losses.

The fraud DNA algorithm is a tool that works with several variables decoded, all packed into five separate layers. These layers are added and the result is called score. Consumer units with score values above 200 are considered suspect, however we need to do a more detailed analysis of the consumption profile and UC features.

Several variables make up DNA algorithm of fraud from information brought by the meter reader, to measurements made feeders and distribution transformers. Each variable is called a predictor. Each has an associated weight, which depicts the level of importance of that variable in the detection of the irregularity.

After all predictors are classified with their respective weights, they are grouped into five distinct layers and each layer is given a weight corresponding to its importance in the final result. As there are no records in the utility of how these weights were found, it can be concluded that many processes may have been used in the determination thereof. This work depicts an alternative that may have been used to obtain the values of the weights of each layer. The hypothesis considered was the use of simple linear regression to determine the best values assigned.

Key words: Commercial losses, DNA fraud, score, decoded, compacted, linear regression

LISTA DE ILUSTRAÇÕES

Figura 2.1 - Entrada de Serviço Monofásica com Pontalete	10
Figura 2.2 – Instalação Correta.....	11
Figura 2.3 - Ligação Adulterada.....	11
Figura 2.4 - Acionamento Remoto	12
Figura 2.5 - Medição Monofásica a Dois Fios, Ligação Correta	13
Figura 2.6 - Medição Adulterada a Dois Fios	14
Figura 2.7 - Composição do DNA da Fraude.....	14
Figura 2.8 - Dispositivo Utilizado para Leitura - PDA	18
Figura 2.9 - Medição de Fronteira na conexão com as Demais Instalações de Transmissão Compartilhadas.....	25
Figura 2.10 - Medição de Fronteira no Ponto de Conexão com a Rede Básica	26
Figura 2.11 - Medição de Fronteira no Ponto de Conexão com a Rede Básica que alimenta um só distribuidor ou consumidor.	26
Figura 2.12 - Medição de Fronteira no Ponto de Conexão com a Rede Básica cuja linha é compartilhada por mais de um distribuidor ou consumidor livre.....	27
Figura 2.13 - Medição de Fronteira das Unidades Geradoras agrupadas por Transformador .	27
Figura 2.14 - Medição de Fronteira das Unidades Geradoras conectadas individualmente....	28
Figura 2.15 - Medição de Fronteira das Unidades Geradoras conectadas através de linhas de transmissão de uso exclusivo.....	28
Figura 2.16 - Medição de Fronteira das Unidades Consumidoras	29
Figura 2.17 - Medição de Fronteira nos pontos de conexão de agentes distribuidores.....	29
Figura 2.18 - Sistema de Medição de Fronteira, concessão Enersul	30
Figura 2.19 – Cadastro dos Transformadores de Distribuição e Perdas Correspondentes.....	31
Figura 2.20 - Cadastro do Transformador, Alimentador Correspondente e Perda do Trafo	32
Figura 2.21 - Pirâmide do Modelo de Segmentação de Fiscalizações	34
Figura 2.22- Matrizes coluna das preditoras	52
Figura 2.23 - Exemplo de utilização do comando zscore.....	52
Figura 2.24- Comando regress determinando peso das Preditoras.....	54
Figura 2.25 - Comando Stepwise	54
Figura 2.26 - Stepwise Regression	55

Figura 2.27 - Tela do comando Stepwise após movimento da preditora X6.....	55
Figura 2.28 - Tela do comando Stepwise após movimento da preditora X7.....	56
Figura 2.28 - Tela do comando Stepwise após movimento da preditora X7.....	56

LISTA DE QUADROS

Quadro 2.1 - Exemplo de linha de código plataforma Qlik View	16
Quadro 2.2 – Composição da Camada N1	17
Quadro 2.3 - Linha de código plataforma Qlik View – Camada N1 – Apontamento do Leiturista.....	21
Quadro 2.4 - Linha de código plataforma Qlik View – Camada N1 – Auto Religado	23
Quadro 2.5 – Linha de código plataforma Qlik View – Camada N1 - Perda Alimentador e Perda CP-Trafo.....	33
Quadro 2.6 - Composição da Camada N2	34
Quadro 2.7 - Linha de código plataforma Qlik View – Camada N2 - Pirâmide	35
Quadro 2.8 - Linha de código plataforma Qlik View – Camada N2 – Variação de Consumo	36
Quadro 2.9 - Linha de código plataforma Qlik View – Camada N2 – Troca de Titularidade	37
Quadro 2.10 - Linha de código plataforma Qlik View – Camada N2 – Inspeções Recentes... ..	38
Quadro 2.11 - Composição da Camada N3.	38
Quadro 2.12 - Linha de código plataforma Qlik View – Camada N3 – Classe Principal	38
Quadro 2.13 - Linha de código plataforma Qlik View – Camada N3 – Tipo de Fase	42
Quadro 2.14 - Composição da Camada N4	43
Quadro 2.15 - Linha de código plataforma Qlik View – Camada N4 – Energia Agregada.....	43
Quadro 2.16 - Linha de código plataforma Qlik View – Camada N4 – Reincidentes	44
Quadro 2.17 - Composição da Camada N5	45
Quadro 2.18 - Linha de código plataforma Qlik View – Camada N5 – Fatura Média	46
Quadro 2.19 – Linha de código plataforma Qlik View – Camada N5 – Contas Vencidas	47
Quadro 2.20 - – Linha de código plataforma Qlik View – Camada N5 – Débito Acumulado	47
Quadro 2.21 - Linha de código plataforma Qlik View – Camada N5 – Classe Principal	47
Quadro 2.22- Linha de código plataforma Qlik View – Camada N5 – Inadimplência.....	49

LISTA DE GRÁFICOS

Gráfico 2.1 – Pesos atribuídos para cada faixa de consumo	22
Gráfico 2.2 – Pesos atribuídos para cada faixa de consumo	35
Gráfico 2.3 - Pesos atribuídos para cada faixa de consumo	40
Gráfico 2.4 - Energia Agregada Per Capita	41
Gráfico 2.5 - Peso atribuído a cada tipo de fase	42
Gráfico 2.6 - Peso atribuído às Classes Principais	48
Gráfico 2.7 - Histórico de Consumo Anual - UC1	58
Gráfico 2.8 - Histórico de Consumo Anual – UC2	59
Gráfico 2.9 - Histórico de Consumo referente aos meses de Março e Abril (2013 e 2014) – UC2.....	60
Gráfico 2.10 - Histórico de Consumo Anual – UC3	61

LISTA DE TABELAS

Tabela 2.1 – Código do Leiturista	19
Tabela 2.2 – Código da Classe Principal Consumidora	39
Tabela 2.3 – Preditoras e Resultado do Parecer	51
Tabela 2.4 – Dados Convertidos em Números	51
Tabela 2.5 – Símbolos das Preditoras Normalizadas	53
Tabela 2.6 – Pesos das Preditoras Determinados por meio do comando Regress	53
Tabela 2.7 - Características UC1	57
Tabela 2.8 - Características UC2.....	59
Tabela 2.9 - Características UC3.....	61

LISTA DE ABREVIATURAS

ANEEL: Agência Nacional de Energia Elétrica

CCEE: Câmara de Comercialização de Energia Elétrica

COFINS: Contribuição para Financiamento da Seguridade Social

CR: Corte Recente

DS: Desligado

DITC: Demais Instalações de Transmissão Compartilhadas

ENERSUL: Empresa Energética de Mato Grosso do Sul

ICMS: Imposto sobre Operações relativas à Circulação de Mercadorias e Prestação de Serviços de Transporte Interestadual e Intermunicipal e de Comunicação

MO: Monofásico

ONS: Operador Nacional do Sistema Elétrico

PDA: Personal Digital Assistant

PIS: Programa de Integração Social

PNT: Perdas Não Técnicas

TOI: Termo de Ocorrência de Irregularidade

TC: Transformador de Corrente

TP: Transformador de Potencial

SUMÁRIO

AGRADECIMENTOS	VI
RESUMO	VIII
ABSTRACT	IX
LISTA DE ILUSTRAÇÕES	X
LISTA DE QUADROS	XII
LISTA DE GRÁFICOS.....	XIII
LISTA DE TABELAS	XIV
LISTA DE ABREVIATURAS	XV
CAPÍTULO 1 - INTRODUÇÃO.....	1
1.1 CONTEXTUALIZAÇÃO	1
1.2 PERDAS COMERCIAIS DE ENERGIA ELÉTRICA.....	3
1.3 OBJETIVOS.....	4
1.4 MOTIVAÇÃO.....	4
CAPÍTULO 2 – FUNDAMENTAÇÃO TEÓRICA.....	6
2.1 Tipos de Perdas Não Técnicas	6
2.1.1 Perdas não técnicas por falha de equipamentos.....	6
2.1.2 Perdas não técnicas devido a erros de faturamento	7
2.1.3 Perdas não técnicas devido à falta de medição	7
2.1.4 Perdas não técnicas devido a anomalias	7
2.1.5 Perdas devido a ação do consumidor.....	8
2.2 Irregularidades no Padrão de Entrada.....	9
2.2.1 Desvio de Energia com Perfuração do eletroduto no ramal de entrada.....	10
2.2.2 Desvio de Energia com Acionamento Remoto.....	12
2.3 Adulteração no Medidor de Energia.....	12
2.4 DNA da Fraude.....	15
2.4.1 Composição do Algoritmo DNA da Fraude	16
2.4.2 Camadas e suas composições	17
2.4.3 Camada N1	17
2.4.4 Apontamento do Leiturista	18
2.4.5 Auto Religado – DS e CR	23

2.4.6 Fator de Perda CP-Trafo e Alimentador - Sistema de Medição de Fronteira e Apuração de Perdas por Subestação, Alimentador e Transformadores de Distribuição.....	25
2.4.7 Camada N2	33
2.4.8 Pirâmide – Estratégia por Segmentação	35
2.4.9 Variação de Consumo.....	36
2.4.10 Troca de Titularidade	37
2.4.11 Inspeções Recentes	37
2.4.12 Camada N3	38
2.4.13 Classe Principal	38
2.4.14 Tipo de Fase.....	40
2.4.15 Camada N4	43
2.4.16 Energia Agregada	43
2.4.17 Reincidentes	44
2.4.18 Camada N5	45
2.4.19 Fatura Média.....	46
2.4.20 Contas Vencidas	46
2.4.21 Débito Acumulado.....	47
2.4.22 Classe Principal	47
2.4.23 Inadimplência	48
2.5 Estrutura do Algoritmo DNA	49
2.5.1 Determinação dos parâmetros do sistema	50
2.5.2 Normalização dos dados.....	51
2.5.3 Determinação dos Pesos das Predictoras.....	53
2.5.4 Conclusões.....	57
2.5.5 Análise dos Resultados	57
CAPÍTULO 3 -CONCLUSÕES	62
CAPÍTULO 4 -REFERÊNCIAS BIBLIOGRÁFICAS	64

CAPÍTULO 1 - INTRODUÇÃO

1.1 CONTEXTUALIZAÇÃO

As perdas não técnicas de energia também conhecidas como perdas comerciais são apuradas pela diferença entre as perdas totais do sistema e as perdas técnicas (inerentes ao transporte da energia na rede elétrica, também conhecidas como perdas fio).

As perdas não técnicas se referem à distribuição e comercialização de energia, sendo consideradas todas as formas de perdas, quer seja por meio da adulteração da medição, erros no processo de medição, faturamentos incorretos, unidades consumidoras sem medição de energia ou erro de cadastro de unidades consumidoras no sistema. [1]

Existem muitos métodos de investigação e detecção de fraudes baseados em sua maioria em técnicas de inteligência artificial, que utilizam informações precisas de bancos de dados, como é o caso da concessionária Enersul. Alguns métodos serão citados a seguir.

Um dos trabalhos realizado por Ghosh e Reilly [2] para um banco americano, *Mellon Bank*, consistiu no uso de redes neurais com a finalidade de detectar fraudes em cartões de crédito da empresa. A rede utilizada foi do tipo radial com três camadas e *feedforward* (alimentação para frente). Para realizar este método, utilizou-se um grande número de transações de cartões de crédito de clientes com as transações já monitoradas, em um espaço de tempo de dois meses. A rede foi treinada utilizando dados de transações honestas e de transações fraudulentas. A rede detectou um grande número de fraudes, sendo que os erros de falsos alarmes foram reduzidos em 20% comparados com a metodologia baseada em regras de análise.

As fraudes na telefonia celular brasileira causam a perda de 5% em sua receita. Taniguchi [3] publicou em 1998 um trabalho que apresentava três propostas de detecção de fraudes para a área de telefonia. O primeiro método utilizava uma rede *feedforward* de inúmeras camadas as quais mapeavam uma função não linear entre duas classes: fraudadores e não fraudadores. Os dados que serviram para treinamento da rede estavam contidos nas faturas dos clientes, os quais retratavam as características dos mesmos. O método pôde ser considerado eficiente já que houve uma taxa de assertividade de 85%

Assim como as instituições financeiras de crédito e telefonia celular as empresas concessionárias também enfrentam sérios problemas de fraudes nos sistemas de faturamento. Consumidores fraudadores com medição em baixa tensão são melhor investigados se forem divididos em classes de consumo. A concessionária já faz esta classificação, que também facilita o controle do faturamento. De acordo com Cabral [4], as unidades consumidoras são classificadas em pequena, média e grande porte, sempre em relação ao consumo de energia. As inspeções em campo são feitas para detecção e combate à continuação das irregularidades, no entanto existe um limite físico do número de fiscalizações e distribuição das equipes de campo. Desta forma, é preciso traçar uma diretriz para busca dos fraudadores. Portanto é criado um perfil de comportamento fraudulento por meio de análises nos bancos de dados das empresas de energia elétrica, gerando regras que classificam os perfis irregulares, para posteriormente traçar um perfil fraudador necessário para centralizar e tornar mais eficiente as fiscalizações futuras. Quando existe um direcionador de inspeções, o número de fiscalizações realizadas é menor, reduzindo custos operacionais e aumentando a produtividade e assertividade do trabalho das equipes de campo.

As considerações de Reis [5] apresentam um sistema que pré seleciona as unidades consumidoras para realizar a inspeção, utilizando o modelo de árvore de decisão. As informações para estudo foram utilizadas do banco de dados de uma distribuidora de energia elétrica, onde 5 fatores foram selecionados e um total de 40.000 clientes foram selecionados. O treinamento ocorreu com a metade dos clientes, que foram escolhidos aleatoriamente, sendo que o teste foi realizado com os restantes. A taxa de assertividade deste procedimento foi de 40%, detectando os fraudadores.

1.2 PERDAS COMERCIAIS DE ENERGIA ELÉTRICA

Um dos grandes desafios enfrentados pelas concessionárias de energia elétrica são as perdas não técnicas. Segundo a agência reguladora de energia elétrica, ANEEL, as perdas comerciais correspondem a parcela de energia não faturada, ou seja todos os tipos de perdas associadas à distribuição de energia elétrica e correspondem a diferença entre as Perdas na Distribuição e as Perdas Técnicas, em megawatt-hora (MWh). Para a definição de perdas comerciais é preciso ter em mente os tipos de perdas do sistema elétrico. São elas: Perdas na Distribuição, Perdas Técnicas e Perdas Não Técnicas.

As perdas na distribuição – PD referem-se a diferença entre a energia injetada na rede da distribuidora e total de energia vendida e entregue, expressa em megawatt-hora e composta pelas perdas de origem técnica e não técnica. [1]

As perdas técnicas representam as perdas inerentes ao processo de transporte da energia, transformação de tensão e medição de energia na rede da concessionária, também expressa em megawatt hora (MWh). [1]

As perdas não técnicas representam todas as demais perdas associadas à distribuição de energia elétrica e correspondem à diferença entre as Perdas na Distribuição e as Perdas Técnicas, em megawatt-hora (MWh). [1]

As perdas comerciais repercutem negativamente para a concessionária e para os consumidores, já que há duas formas de conter os prejuízos gerados: revisão tarifária, prevista pela ANEEL, a qual permite que a concessionária repasse ao consumidor uma parcela das perdas e a outra é quando a concessionária assume o percentual não repassado na tarifa, gerando prejuízo aos acionistas e tornando o serviço deficitário, já que os investimentos que deveriam ser destinados a melhoria do atendimento devem cobrir os prejuízos.

As perdas não técnicas podem ser classificadas de diversas formas, no entanto as irregularidades causadas por agentes externos causam maior prejuízo às concessionárias de energia elétrica.

Alguns aspectos negativos acerca das práticas irregulares no processo de comercialização de energia estão descritos a seguir:

- Risco de Acidentes: Muitas pessoas se arriscam na realização de fraudes, comprometendo sua própria segurança e muitas vezes a segurança das pessoas a sua volta.
- Evasão de Tributos como: ICMS, PIS COFINS e iluminação pública. As fraudes de energia elétrica resultaram em prejuízo de R\$ 32 milhões aos cofres públicos de Mato Grosso do Sul em 2013.¹
- Concorrência abusiva no mercado, a medida que o consumidor que comete a irregularidade tem a possibilidade de vender seus produtos a um preço menor em relação aos demais concorrentes que arcam com suas despesas.
- Aumento da tarifa, que reduz as possibilidades de crescimento de mercado e abertura de novos empreendimentos.
- Diminuição do faturamento das concessionárias e uma consequente queda nos investimentos e melhorias no sistema elétrico.
- Comprometimento dos indicadores de qualidade das concessionárias, já que muitas irregularidades provocam desligamentos simultâneos e imprevistos na rede elétrica.

¹ Dados estimados em estudo realizado pelo Departamento de Medição e Combate às Perdas Comerciais da Enersul, referentes ao ano de 2013.

1.3 OBJETIVOS

Estudar o modelo empregado pela concessionária de energia, o qual após ser adotado reduziu o valor das perdas comerciais de energia elétrica, o número de fiscalizações e aumentou a taxa de assertividade das equipes de campo, otimizando o trabalho da concessionária, no que se refere ao combate de irregularidades e realização das fiscalizações.

1.4 MOTIVAÇÃO

Devido ao alto índice de unidades consumidoras com situações irregulares, a concessionária de energia elétrica precisou buscar soluções para combater às irregularidades de energia elétrica de maneira assertiva e eficiente. O uso da ferramenta DNA da fraude aumentou a taxa de sucesso das fiscalizações, reduzindo os gastos da concessionária, aumentando o valor da energia agregada e gerando menos perdas aos cofres da empresa.

É importante que as perdas comerciais sejam reduzidas, já que todos os consumidores assumem uma parcela dos dividendos subtraídos da receita, portanto é preciso reduzir cada vez mais a parcela particionada aos consumidores honestos, que assumem uma dívida, a qual não tem responsabilidade.

A concessionária de energia elétrica tem grande interesse na redução das perdas não técnicas, afinal a sua credibilidade com os acionistas é recuperada e sua receita é preservada, possibilitando o aumento dos investimentos em outras áreas da distribuição de energia, aumentando a confiabilidade do sistema e gerando satisfação para os consumidores.

CAPÍTULO 2 - FUNDAMENTAÇÃO TEÓRICA

2.1 Tipos de Perdas Não Técnicas

As perdas comerciais podem se originar de problemas na medição classificados como falhas na medição, irregularidades técnicas ou irregularidades na medição. As falhas na medição são todos os problemas gerados por defeitos inerentes ao sistema de medição, que não foram causados intencionalmente pelo consumidor.

A irregularidade técnica é todo o tipo de adulteração no sistema de medição e que não é feita diretamente no medidor. Estão configuradas e enquadradas nesse grupo os desvios de energia pelo ramal de entrada², ligações clandestinas e desvio de energia antes da medição no interior da mureta.

A irregularidade na medição é toda adulteração realizada no medidor. Geralmente, o circuito eletrônico do medidor é adulterado, havendo casos recentes, onde medidores eletrônicos foram reprogramados tendo suas constantes de medição alteradas, reduzindo o faturamento da unidade consumidora.

2.1.1 Perdas não técnicas por falha de equipamentos

Os equipamentos de medição de energia elétrica estão sujeitos a falhas, sendo impraticável a modelagem do grau de deterioração destes equipamentos. O medidor eletromecânico pode sofrer influências de elementos externos, como a variação da temperatura ambiente, variações da tensão da rede, da frequência, influência de campos magnéticos externos e distorções harmônicas [6]. As ligações clandestinas feitas nos equipamentos de medição, em sua maioria medidores de energia elétrica, utilizam condutores com bitola inferior à recomendada, agravando o fator perdas e não levando em consideração as normas técnicas de segurança e confiabilidade do sistema.

² Conjunto de condutores, equipamentos e acessórios que interligam o ponto de entrega da energia ao ponto de proteção, medição ou transformação, inclusive.

2.1.2 Perdas não técnicas devido a erros de faturamento

Os erros de faturamento também compõem o montante das perdas comerciais. Quando ocorre a leitura na unidade consumidora e esta é feita de forma equivocada pelo leiturista, ocorre um erro no faturamento e ocorrendo um acréscimo da perda não técnica, já que parte da energia que deveria ter sido faturada se perdeu no processo da leitura. A compensação desta perda é feita no ciclo de faturamento posterior à leitura errônea.

Juntamente aos erros de medição e faturamento estão os erros do sistema comercial de faturamento, os quais podem gerar valores errados, troca de faturas entre clientes e problemas na contabilidade do total da energia faturada.

2.1.3 Perdas não técnicas devido à falta de medição

A falta de medidores de energia em uma unidade consumidora pode ser um erro da concessionária de energia, que ligou, provisoriamente, uma UC sem o medidor e não concretizou sua ligação definitiva. Outro tipo de ausência na medição está definida na legislação, para consumidores sem medição, como é o caso de bancas de jornal, quiosques e iluminação pública. Para esses tipos de consumidores existe o consumo estimado, os quais podem contribuir para mais ou para menos no cálculo das PNT.

2.1.4 Perdas não técnicas devido a anomalias

Os medidores de energia elétrica podem apresentar anomalias em seu funcionamento, devido a problemas intrínsecos ao seu longo período de funcionamento e às suas condições de uso. Alguns problemas na estrutura do medidor não causam alteração na medição de energia e no faturamento, como é o caso do vidro do mostrador quebrado. Existem algumas alterações nos medidores que diminuem significativamente seu poder de medição, como é o caso de queima de bobinas devido a descargas atmosféricas, disco arranhado ou queima do circuito eletrônico.

2.1.5 Perdas devido a ação do consumidor

As perdas causadas pela ação do consumidor podem ser definidas como fraudes ou furtos de energia, sendo que existem diferenças sutis entre elas. A fraude de energia elétrica é uma atitude intencional de reduzir o valor da energia consumida, sem pagar pelo valor de direito.

O furto é configurado sempre que ocorre ligações direto da rede e que não tem a presença do medidor de energia elétrica, geralmente ocorrem em regiões de conflito, onde ocorrem invasões de propriedades e ligações clandestinas.

A fraude é uma adulteração na medição, totalitariamente ou parcialmente. Após a adulteração da medição é caracterizado um aumento do consumo por parte da unidade consumidora. Algumas irregularidades foram caracterizadas pela agência reguladora, ANEEL, após sucessivas análises dos perfis fraudadores. São elas:

- Ligação direta com a rede da concessionária;
- Desvio da medição antes do ramal de entrada (desvio na mureta);
- Medidor com a carcaça furada com objeto estranho inserido em seu interior, dificultando ou interrompendo a movimentação do disco do medidor, alterando a medição;
- Ponteiros deslocados propositalmente;
- Medidor com ligação dos seus terminais de corrente ou tensão invertidas;
- Terminal de prova aberto;
- Bobina de potencial interrompida;
- Engrenagem do medidor adulterada;

- Elemento móvel empenado;
- Chave de Aferição Aberta;
- Condutores entre a chave de aferição e medidor isolados;
- Condutores seccionados entre a chave de aferição e medidor;
- Sequência de fase adulterada e
- Curto circuito provocado na entrada e saída do medidor.

2.2 Irregularidades no Padrão de Entrada

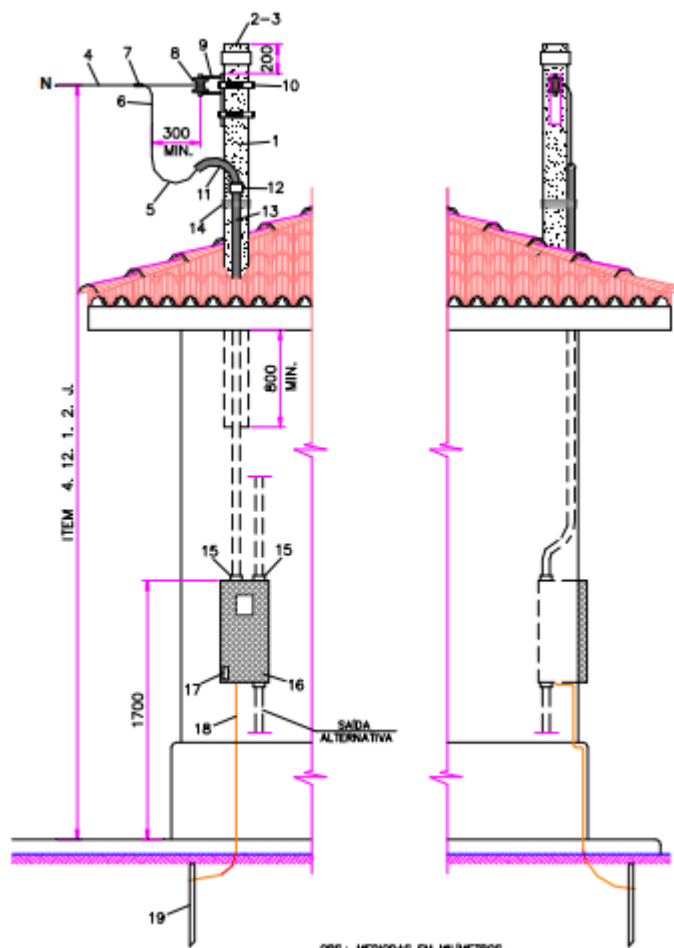
Padrão de entrada é o conjunto de instalações composto de caixa de medição, sistema de aterramento, condutores e acessórios indispensáveis para realização da ligação da unidade consumidora. Os padrões de entrada de serviço deverão ser escolhidos de acordo com as características das cargas instaladas, as quais determinam o tipo de fornecimento.³

Cabe ao consumidor preparar toda a entrada de serviço, inclusive o padrão de entrada próprio. Portanto, cada consumidor é classificado quanto ao tipo de fornecimento de energia elétrica para atendimento em tensão secundária. Qualquer aumento de carga ou alteração de suas características, além do limite previsto em norma da concessionária deve ser previamente submetido à apreciação da Enersul, sob pena de suspensão no fornecimento de energia caso ocorra algum tipo de prejuízo no atendimento às outras unidades.

A figura 2.1 indica uma entrada típica de serviço monofásica com pontalete. Neste caso a instalação da caixa de medição deve ser feita na parede.

³ Norma de Distribuição – TDE – 102. Fornecimento de Energia Elétrica em Tensão de Distribuição Secundária – 220/127. Enersul, 11/09/2013.

Figura 2.1 - Entrada de Serviço Monofásica com Pontalete



Fonte: Norma de Distribuição 102 - Enersul

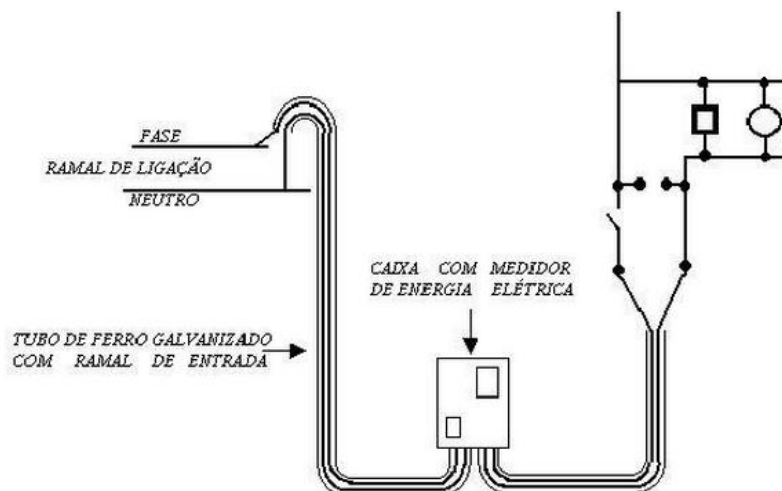
2014

2.2.1 Desvio de Energia com Perfuração do eletroduto no ramal de entrada

Trata-se de uma irregularidade que não é feita no medidor, mas sim um desvio de energia feito por meio de uma perfuração no eletroduto galvanizado do ramal de entrada. Esse desvio normalmente é executado no interior da parede, muro ou mureta.

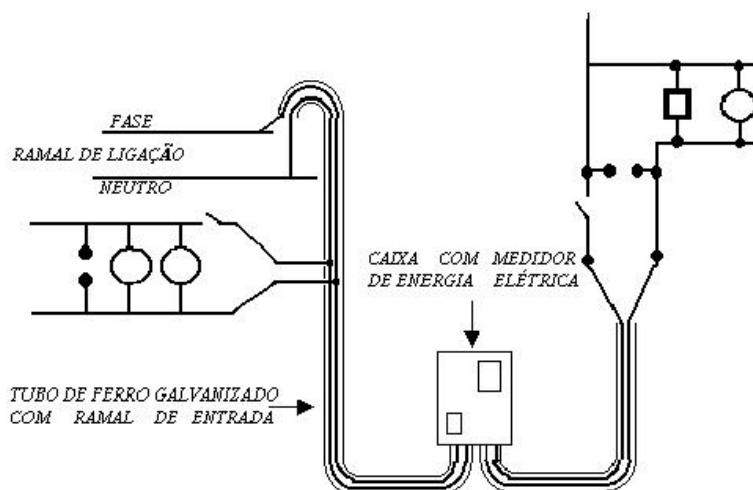
Parte da corrente que atravessa os condutores é interceptada e desviada para a carga, não sendo computada na medição.

Figura 2.2 – Instalação Correta



Fonte: Fraudes no Padrão de Entrada⁴

Figura 2.3 - Ligação Adulterada



Fonte: Fraudes no Padrão de Entrada³

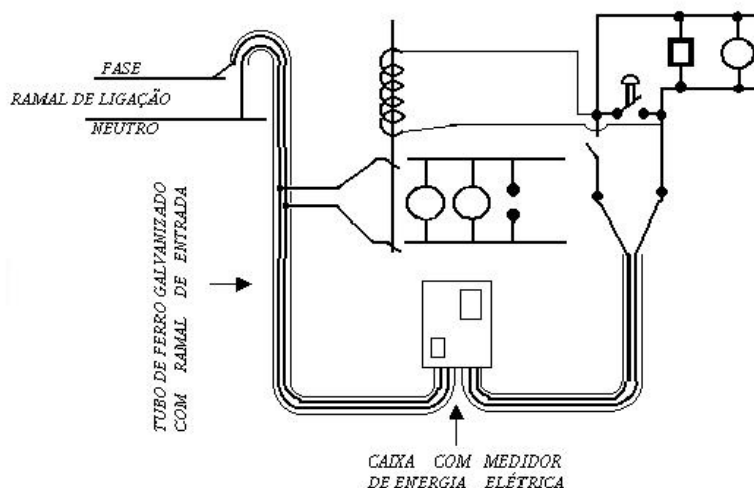
⁴ Disponível em <https://sites.google.com/site/punarobley/fraudes-mais-comuns-na-medicao-de-energia-eletrica>. Acesso em dezembro de 2014.

2.2.2 Desvio de Energia com Acionamento Remoto

O desvio de energia por acionamento remoto é feito por meio de perfuração no eletroduto. A grande diferença entre esta irregularidade e a anterior é que existe um controle da corrente que chega ao medidor.

Com o acionamento remoto é mais difícil de detectar a fraude, pois o fraudador controla os momentos em que o controle atuará. No momento de uma inspeção, por exemplo é desligado o acionamento e, portanto toda a energia requerida pelo sistema está sendo medida.

Figura 2.4 - Acionamento Remoto



Fonte: Fraudes no Padrão de Entrada⁵

2.3 Adulteração no Medidor de Energia

A adulteração no medidor pode ser provocada de diversas formas, tais como retirada da tampa de vidro, com violação dos lacres de aferição. As seguintes ações são efetuadas:

- Manipulação da posição dos ponteiros;

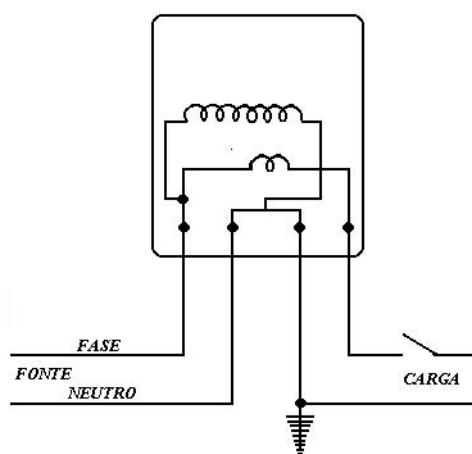
⁵ Disponível em <https://sites.google.com/site/punarobley/fraudes-mais-comuns-na-medicao-de-energia-eletrica>. Acesso em dezembro de 2014.

- Introdução de objetos estranhos no interior do disco, fazendo com que seu movimento seja alterado;
- Manipulação da fiação interna;
- Adulteração nas bobinas de corrente e/ou tensão.

Os fraudadores preferem corromper os medidores sem retirar a tampa de vidro. As técnicas utilizadas são:

- Perfuração da carcaça ou da tampa e introdução de objetos, que alteram os mecanismos internos do medidor;
- Manipulação da fiação de ligação do medidor invertendo o sentido da corrente nas bobinas do medidor (figura 2.6);

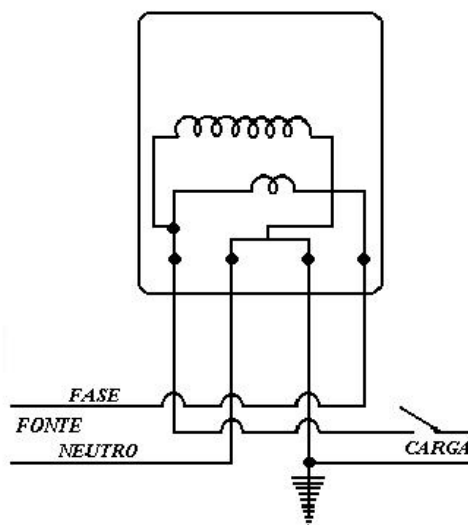
Figura 2.5 - Medição Monofásica a Dois Fios, Ligação Correta



Fonte: Fraudes no Padrão de Entrada⁶

⁶ Disponível em <https://sites.google.com/site/punarobley/fraudes-mais-comuns-na-medicao-de-energia-eletrica>. Acesso em dezembro de 2014.

Figura 2.6 - Medição Adulterada a Dois Fios



Fonte: Fraudes no Padrão de Entrada⁷

O medidor eletromecânico possui dois tipos de bobinas, de corrente e de tensão. A figura 2.5 indica uma ligação correta das bobinas de corrente, que possuem apenas duas espiras e das bobinas de tensão, com número maior de espiras.

Um dos tipos mais comuns de irregularidades é a inversão do sentido de corrente das bobinas, como é mostrado na figura 2.6. A alimentação da rede deve ser conectada ao terminal positivo da bobina e o terminal negativo deve ser ligado à carga. Em uma ligação adulterada, a rede é ligada no terminal negativo da bobina, enquanto que a carga é ligada no terminal positivo.

A medição fica comprometida com a inversão das ligações nas bobinas de corrente, reduzindo o valor medidor.

⁷ Disponível em <https://sites.google.com/site/punarobley/fraudes-mais-comuns-na-medicao-de-energia-eletrica>. Acesso em dezembro de 2014.

2.4 DNA da Fraude

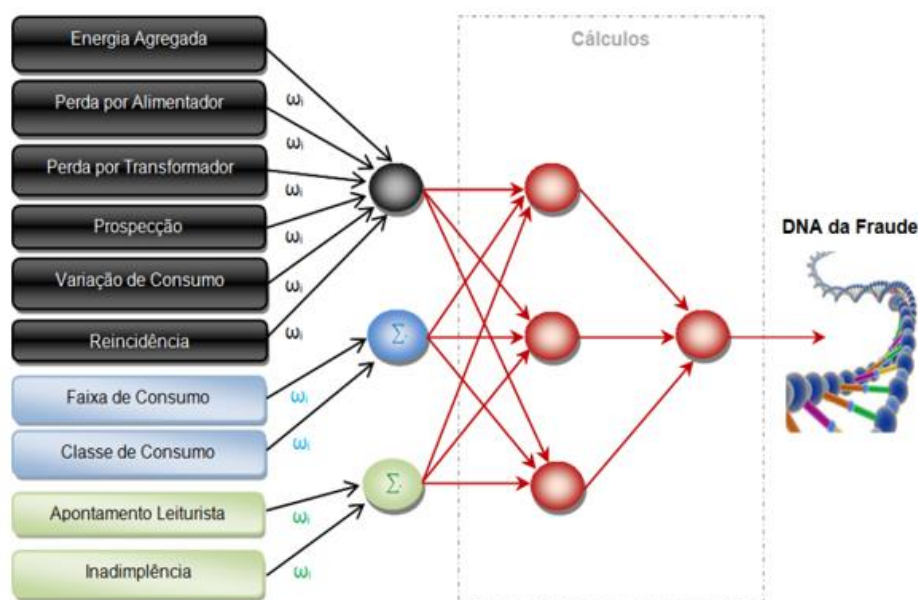
O DNA da Fraude é um indicador da probabilidade de anomalias na medição, que tem por objetivo elevar a acurácia e a assertividade na identificação de fraudes e otimização da energia agregada.⁸

A metodologia desta ferramenta utiliza a ideia de uma rede composta por variáveis decodificadas, onde para cada parcela é creditado um peso, dependendo de sua complexidade e importância para a saída do sistema.

O DNA é um algoritmo que utiliza variáveis decodificadas fazendo sua compactação, ou seja reduz um grande número de variáveis em apenas 5, traduzindo de forma simplificada seu conteúdo.

O algoritmo do DNA da Fraude foi implementado na plataforma QlikView ® (*Business Intelligence*).

Figura 2.7 - Composição do DNA da Fraude



Fonte: Enersul (2014)

⁸ O termo energia agregada refere-se ao montante de energia, que é inserida no sistema após a regularização da unidade consumidora, ou seja o montante de energia que torna a ser fatura.

2.4.1 Composição do Algoritmo DNA da Fraude

O algoritmo do DNA é composto por diversas variáveis, agrupadas em camadas, num total de cinco, dispostas paralelamente. Cada variável possui um nível de importância para a composição do DNA. Os pesos foram definidos de acordo com o nível de impacto que aquela variável tem sobre a descoberta da irregularidade e foram determinados por meio de análise Heurística⁹.

Todo o processo de decodificação do algoritmo é mostrado a seguir. Os dados são compactados e reajustados em cinco camadas. Ou seja, inúmeras variáveis se transformam em apenas cinco, traduzindo o significado do algoritmo simplificadamente.

O resultado do DNA é um somatório de valores decodificados e o seu valor é denominado score. Após a unidade consumidora ter todas as suas variáveis somadas e com seus respectivos pesos atribuídos, o score encontrado pode definir o perfil da UC em questão, podendo ela ser ou não considerada como potencial fraudadora.

A linha de comandos é ilustrada no quadro 2.1. As camadas **N1, N2, N3, N4 e N5** compõem o cálculo, sendo que cada uma delas é composta por várias variáveis, as quais recebem pesos específicos de acordo com sua importância e seu impacto na determinação e na assertividade da irregularidade.

Quadro 2.1 - Exemplo de linha de código plataforma Qlik View

if(<i>Linha 1</i>
(<i>Linha 2</i>
2*\$(N1) + \$(N2) + \$(N3) + \$(N4) + \$(N5))>200,	<i>Linha 3</i>
(<i>Linha 4</i>
2*\$(N1) + \$(N2) + \$(N3) + \$(N4) + \$(N5))	<i>Linha 5</i>
, ' ')	<i>Linha 6</i>

Fonte: Mariana Viudes Villalba (2014)

⁹ É um procedimento que encontra soluções de uma maneira fácil utilizando a experiência, a intuição e o conhecimento empírico.

O significado das linhas mais importantes é descrito a seguir:

Linha 3 – Soma todas as camadas e atribui uma condição para o valor resultante, no caso 200. Ou seja, o resultado pode ser maior ou menor que 200.

Linha 5 – Se o valor encontrado anteriormente for maior que duzentos o valor atribuído ao score é a soma das camadas.

Linha 6 – Caso o resultado da operação seja menor que 200, nenhum valor é atribuído ao score.

2.4.2 Camadas e suas composições

Cada camada do algoritmo é composta por diversas variáveis, sendo que para cada variável é atribuído um peso específico.

2.4.3 Camada N1

A camada N1 é composta por quatro variáveis: Apontamento do leiturista, cliente com situação Auto religado com corte recente ou com desligamento no sistema, Fator de Perda por alimentador e Fator de Perda CP-Trafo.

A sequência de comandos da camada N1 está detalhada no quadro 2.2.

Quadro 2.2 – Composição da Camada N1

//Apontamento do Leiturista
+
//Auto Religado – DS e CR
+
//Fator de Perda Alimentador
+
//Fator de Perda CP-Trafo

Fonte: Elaborado pelo autor (2014)

Todas as variáveis são somadas para a composição da camada N1. É certo, a maioria dos casos de auto religação foram constatados pelo leiturista, por isso a união destas duas variáveis na composição da mesma camada.

As variáveis fator de perda alimentador e fator de perda CP-Trafo estão diretamente relacionadas, já que o resultado da perda no transformador (CP-Trafo) reflete o resultado da perda no alimentador.

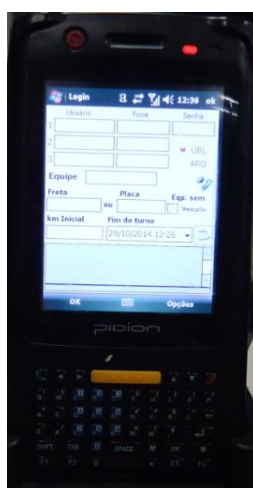
É questionável a composição de uma única camada abranger as quatro variáveis, já que juntas, elas não são correlacionadas¹⁰. Um ponto de melhoria para o processo seria a decomposição destas camadas em outras duas, onde uma delas agruparia as variáveis Apontamento do Leiturista e Auto religado e a outra camada seria composta pelos fatores Perda Alimentador e Perda CP-Trafo.

2.4.4 Apontamento do Leiturista

O faturamento da maioria das unidades consumidoras de Mato Grosso do Sul é feito por intermédio de leitura visual, que é feita com o auxílio de um profissional leiturista. Não é de responsabilidade do leiturista detectar irregularidades na medição, no entanto os profissionais da leitura podem contribuir para a detecção de anormalidades.

Para cada situação que possa caracterizar a fraude, o leiturista faz um tipo de anotação em um dispositivo chamado PDA. Nesse aparelho, existem códigos pré definidos, que caracterizam todas as possíveis medições duvidosas encontradas em campo.

Figura 2.8 - Dispositivo Utilizado para Leitura - PDA



Fonte: Enersul (2014)

¹⁰ Diz-se que existe correlação entre variáveis quando as alterações sofridas por uma delas são acompanhadas por modificação nas outras.

Os códigos apresentados pelo leiturista estão descritos na tabela 2.1.

Tabela 2.1 - Código do Leiturista

Código Do Leiturista	Apontamento	Peso Atribuído
004	Medidor com Defeito (Área Urbana)	10
031	Constante do Medidor Incorreta	10
104	Medidor com Defeito (Área Rural)	10
314	Medidor Girando ao Contrário	10
033	Faturamento feito por média (Trafo/ TC/ TP danificados)	6
037	Medidor Eletrônico com Led Apagado	6
034	Suspeita de Fraude	5
012	Inspecionar Entrada de Serviço	2
029	Ponteiro Deslocado	2
015	Led Apagado	1

Fonte: Elaborado pelo autor (2014)

O leiturista tem autonomia para adicionar qualquer código de leitura e vinculá-lo à unidade consumidora. As irregularidades mais comuns estão descritas na Tabela 2.1.

Medidores com defeito, tanto em áreas urbanas como em áreas rurais recebem o código 004 e 104. Estes códigos caracterizam os medidores que não marcam a leitura corretamente, impossibilitando sua anotação.

Todos os medidores apresentam constantes de multiplicação, que dependem do tipo de medição, das características dos medidores e das relações de transformação de corrente e de potencial. Nas medições diretas (grupo B), o fator de multiplicação caracteriza as propriedades do registrador do medidor, já que estes registram a totalidade da energia consumida [7]. Caso haja divergência da constante de cadastro do medidor, no sistema, com a constante encontrada em campo, o leiturista aponta em seu dispositivo o erro da constante de multiplicação do medidor e atribui o código 031 para aquela UC.

Outro fenômeno bastante recorrente é a adulteração dos ponteiros existentes nos medidores eletromecânicos, conhecidos como disco, que podem eventualmente girar ao contrário, quer seja por defeito próprio do equipamento, quer seja por meio de manipulação de agente externo. Caso o leiturista aponte tal fenômeno, ele anota o código 314.

A resolução 414/2010 da Aneel refere-se às unidades consumidoras com fornecimento de energia elétrica atendidas em alta tensão (13,8 kV ou 34,5 kV), que podem optar por outro faturamento.

. Para esse tipo de enquadramento a unidade deve ser de pequeno porte, com potência instalada em transformadores igual ou inferior a 112,5 kV e realizar contrato de fornecimento [8]. A medição de unidades consumidoras grupo A optantes é composta de transformadores de potencial e transformadores de corrente, denominados TP's e TC's, respectivamente. Caso estes instrumentos estejam danificados, a unidade consumidora é faturada pela média dos doze últimos faturamentos. O leiturista aponta as deficiências dos equipamentos de medição e o faturamento por média, utilizando o código 033.

Os medidores eletrônicos possuem leds que indicam a medição nas três fases correspondentes. Caso os leds estejam apagados, é possível concluir que uma das fases não está registrando consumo. Se o leiturista apontar tal deficiência o código anotado é 037 ou 015.

Outro indicativo que gera o apontamento do leiturista são os lacres violados do medidor e a entrada de serviço com suspeita de desvio de uma das fases na mureta. O desvio de energia é uma das irregularidades mais difíceis de ser atestadas, já que a fraude é feita no interior da parede. Caso o leiturista ateste qualquer irregularidade nos lacres ou na mureta ele aponta os códigos 034 e 012, para suspeita de fraude devido violação dos lacres e suspeita de desvio na mureta, respectivamente.

Para o caso de ponteiro deslocado, no caso de medidores eletromecânicos existe também um código de apontamento – 029.

O último código é para o sistema de tele medição. Algumas unidades consumidoras rurais não recebem a visita do profissional leiturista, elas são medidas remotamente. Caso o faturamento dessas unidades esteja sendo feito pela média, ou não esteja sendo feito corretamente por qualquer outro motivo, um leiturista visita o local da tele medição pessoalmente e atesta se há algum problema de comunicação. Este código é o 032.

Todas estas anotações de leitura entram na composição da camada N1 e cada uma delas possui um peso característico, que depende de sua importância na definição da camada. Experimentalmente, foi atestado que os códigos que apresentam maior indicativo de fraude e maior acerto por parte do leiturista foram os códigos: 004 e 104 (medidor com defeito – urbano e rural), 031 (constante do medidor incorreta) e 314 (medidor girando ao contrário). Para estes códigos foi definido o peso 10.

Os outros dois códigos mais importantes foram os códigos 033 (faturamento por média) e 037 (medidor eletrônico com led apagado), que receberam peso 6. Em menor escala estão os códigos 034 (suspeita de fraude), 012 (inspecionar entrada de serviço) e 029 (ponteiro deslocado), que recebem o peso 2.

Caso o apontamento feito seja 015 (led apagado – medidor eletromecânico), o peso atribuído será 1. O menor peso foi atribuído a esse apontamento, pois a taxa de assertividade nele é bem pequena e o led apagado do medidor eletromecânico, na maioria das vezes não configura fraude, podendo ser apenas um defeito de visualização da leitura, não interferindo de forma significativa para a queda do faturamento.

A linha de código da variável apontamento do leiturista está descrita a seguir no quadro 2.3.

Quadro 2.3 - Linha de código plataforma Qlik View – Camada N1 – Apontamento do Leiturista

```
// Apontamento do Leiturista
40 * ((                                // Valor supostamente encontrado por meio de regressão linear simples
1+                                     // Soma-se 1 à linha de código, para evitar que o resultado do algoritmo se anule
if (Cod_Mensagem_Leiturista_Apontamento='004' or      // Se o código do leiturista for 004 ou
Cod_Mensagem_Leiturista_Apontamento = '031' or      // se o código for 031 ou
Cod_Mensagem_Leiturista_Apontamento = '104' or      // se o código for 104 ou
Cod_Mensagem_Leiturista_Apontamento = '314', 10,      //se o código for 314, atribui-se peso 10
if (Cod_Mensagem_Leiturista_Apontamento='033' or      // Senão, se o código for 033 ou
Cod_Mensagem_Leiturista_Apontamento = '037', 6,      // se o código for 037, atribui-se peso 6
if (Cod_Mensagem_Leiturista_Apontamento='034', 5,      // Senão, se o código for 034, atribui-se peso 5
if (Cod_Mensagem_Leiturista_Apontamento='012' or      // Senão, se o código for 012 ou
Cod_Mensagem_Leiturista_Apontamento = '029', 2,      // se o código for 029, atribui-se peso 2
if (Cod_Mensagem_Leiturista_Apontamento='015', 1))))), // Senão, se o código for 015, atribui-se peso 1
0)) // Caso não satisfaça nenhuma das condições anteriores, nenhum valor é atribuído à variável
```

Fonte: Elaborado pelo autor (2014)

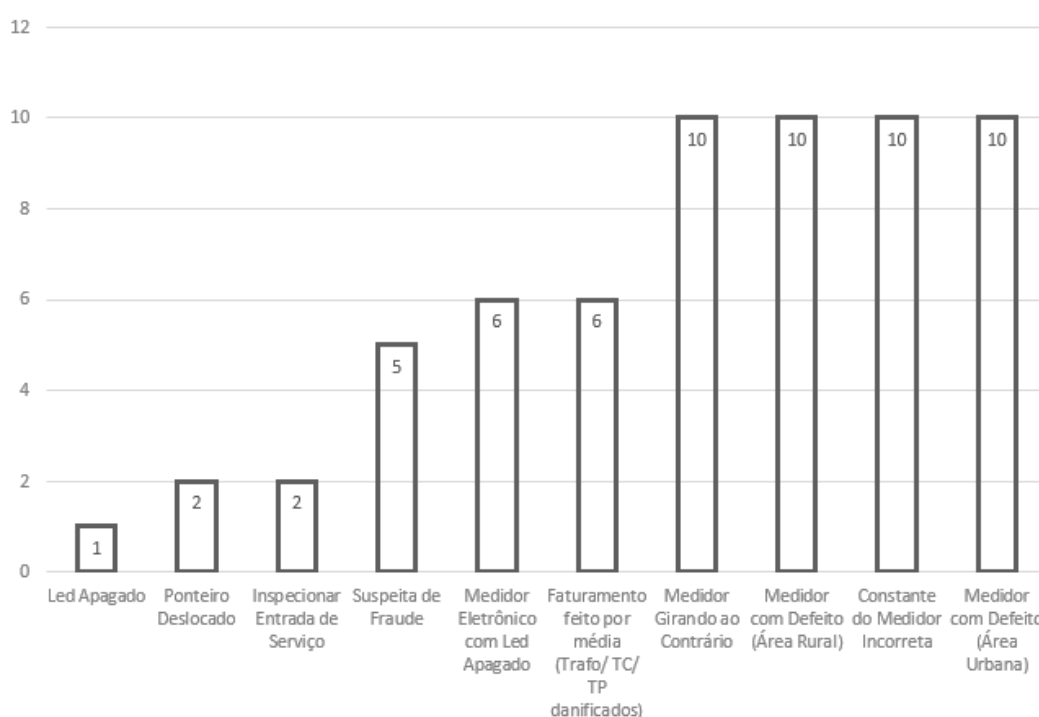
Como não há registros nos arquivos da empresa Enersul a respeito da metodologia empregada na determinação dos pesos e do valor 40 utilizado para composição da variável, é possível que várias ferramentas podem ter sido utilizadas na determinação desses valores como PCA (principal componente analysis), Backprogramation, Heurística ou Regressão Linear.

Este trabalho partirá do pressuposto que os pesos de cada variável foram determinados Heuristicamente, já que os valores são inteiros e bem definidos, no entanto não se pode descartar a possibilidade destes pesos terem sido determinados por métodos estatísticos. Com a experiência das equipes de campo e do operar do DNA, é possível determinar o nível de importância de cada variável para a atribuição dos respectivos pesos. A empresa apresenta os históricos de apontamentos que trouxeram resultados positivos em relação à detecção da fraude. Com estes dados é possível selecionar quais apontamentos tiveram melhor desempenho. Assim a previsão dos pesos pode ser baseada em comportamentos do passado.

Em relação ao valor quarenta é possível que tenha sido utilizado um método de regressão linear simples para determinação do melhor valor que representa a camada N1. Para a determinação deste valor é preciso uma amostra aleatória de unidades consumidoras, que já tenham sido fiscalizadas com os resultados das fiscalizações conhecidos e que tenham em seus registros os apontamentos do leiturista, se estavam ou não auto religadas, o valor das perdas no alimentador e trafo correspondentes.

. O nível de importância de cada peso está ilustrado no gráfico 2.1

Gráfico 2.1 - Pesos atribuídos aos Apontamentos do Leiturista



Fonte: Elaborado pelo autor (2014)

2.4.5 Auto Religado – DS e CR

Uma característica típica dos fraudadores, que estão desligados no sistema (DS) ou que estão com corte recente (CR) é a auto religação da energia, que pode ser feita no próprio medidor ou direto da rede de distribuição. Nestes casos, a Enersul não recebe pelo consumo, já que no sistema o cliente está desligado. A composição da variável auto religação associa a situação desligado ou cortado com dois apontamentos feitos pelo leiturista. Isso significa que existe uma relação entre as situações CR ou DS e os códigos 034 e 314, ou seja em muitas autuações o cliente desligado ou cortado com auto religação, teve apontamento de leitura com os códigos 034 e 314. Estes códigos referem-se a suspeita de fraude e medidor girando ao contrário, respectivamente.

A ferramenta Qlick View importa dados externos vindos da base de dados da Enersul. Os dados vêm dos mais diversos setores como faturamento e leitura. Um desses dados refere-se aos apontamentos feitos pelo leiturista. As informações disponíveis são: unidade consumidora apontada, a suspeita registrada pelo apontamento e a data em que foi registrado o apontamento. Por meio dessas informações é possível compor o código da variável Auto Religado.

A variável Ref_Apontamento foi criada para agrupar duas informações referentes ao mês e ano da geração do apontamento. A variável Data_Energia_Agregada contém as informações da data atual da energia agregada que está sendo calculada, ou seja refere-se ao mês e ano atuais. Sendo assim, é possível relacionar a data da referência do apontamento e a data atual em uma mesma variável, no caso Auto Religado.

Quadro 2.4 - Linha de código plataforma Qlik View – Camada N1 – Auto Religado

```
// Auto Religado – DS e CR
if ((                                     // se
(CodSitUC = 'DS' or CodSitUC = 'CR' and   // a UC está desligada ou cortada do sistema e
Ref_Apontamento = '$(Data_Energia_Agregada)' and // o mês de apontamento corresponde ao mês atual
(Cod_Mensagem_Leiturista_Apontamento = '034' or // e o código do leiturista for suspeita de fraude ou
Cod_Mensagem_Leiturista_Apontamento = '314')), // o código for medidor girando ao contrário
10, 1) // e se todas as condições forem satisfeitas atribui-se o valor 10, senão atribui-se o valor 1
```


2.4.6 Fator de Perda CP-Trafo e Alimentador - Sistema de Medição de Fronteira e Apuração de Perdas por Subestação, Alimentador e Transformadores de Distribuição.

O processo de medição de fronteira é utilizado para monitorar a energia injetada e absorvida no sistema elétrico. Para ser feita a contabilização da Câmara de Comercialização de Energia Elétrica, que calcula os encargos do uso dos sistemas de transmissão e verifica as capacidades declaradas de geração e o cumprimento das instruções de despacho. As medições de faturamento de fronteira devem seguir um critério de instalação dos medidores nos seguintes pontos [9]:

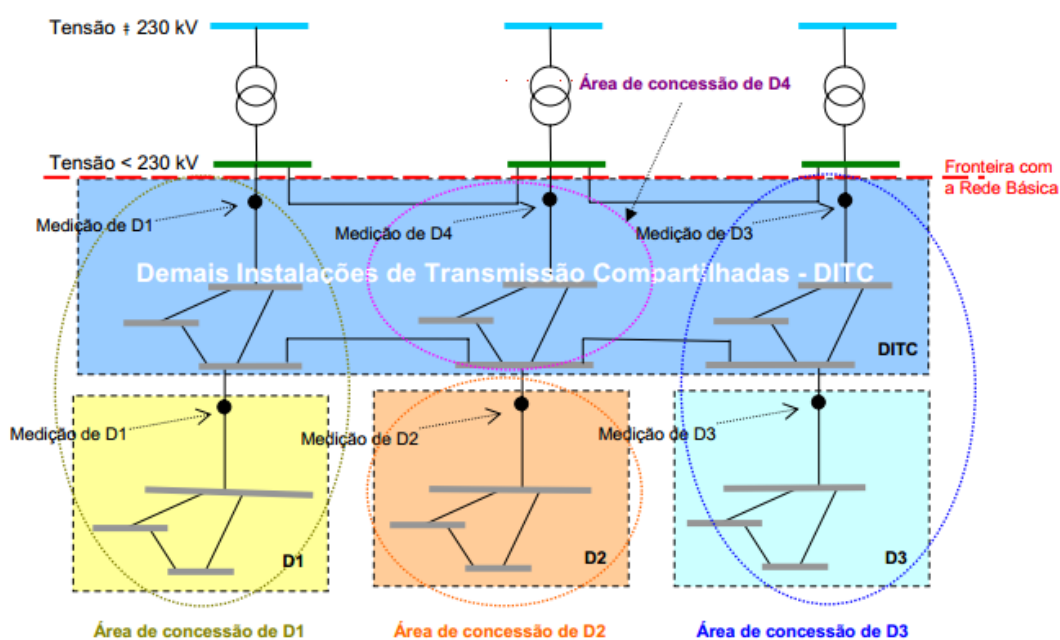
- Na conexão com a rede básica;
- Na conexão com as demais instalações de Transmissão Compartilhadas – DITC;
- Na conexão de consumidores livres;
- Nas unidades geradoras onde existe contabilização de serviços ancilares.¹¹
- Na conexão entre sistemas de agentes que fazem parte da CCEE;
- Na interligação internacional (importação ou exportação de energia) nos sistemas interligados;
- Na interligação entre sub mercados;
- Nas unidades geradoras das usinas despachadas centralizadamente pelo ONS, para medição de geração bruta;

¹¹ Serviços ancilares são serviços prestados por unidades geradoras que proveem energia reativa ao sistema, sendo que estas unidades podem operar como compensador síncrono, compensador automático de controle de geração (CAG), auto restabelecimento (Black start) ou sistemas especiais de proteção.

- Nas unidades geradoras ou por grupo de unidades geradoras, para a medição de geração líquida;
- Na conexão de auto produtor;
- Serviços auxiliares de usinas e subestações.

A localização dos pontos de medição, estão ilustradas nas figuras a seguir:

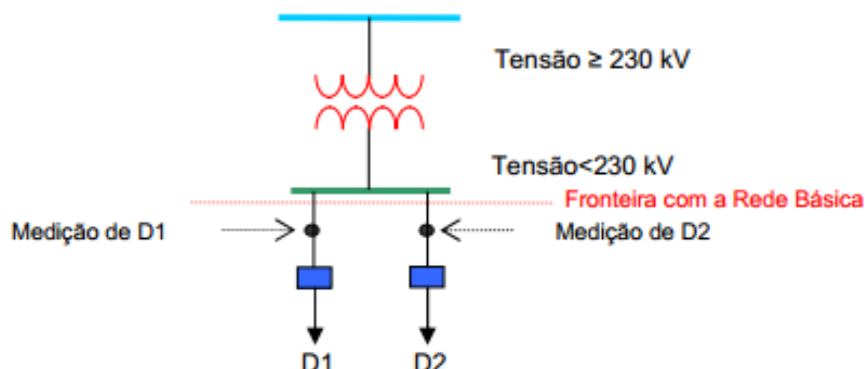
Figura 2.9 - Medição de Fronteira na conexão com as Demais Instalações de Transmissão Compartilhadas



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

No ponto de conexão com a rede básica que alimenta mais de um distribuidor ou consumidor, a medição deve ser individual por alimentador no lado de baixa tensão no transformador.

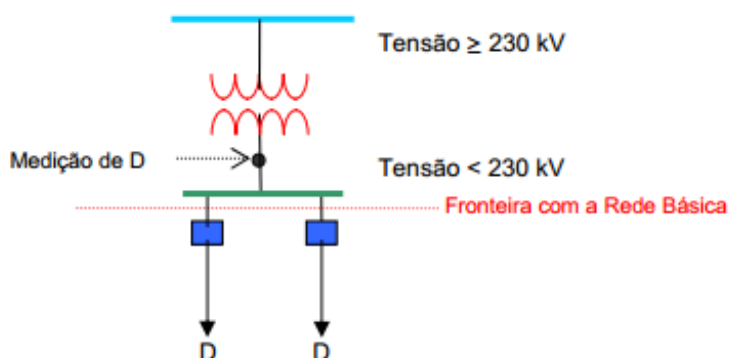
Figura 2.10 - Medição de Fronteira no Ponto de Conexão com a Rede Básica



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

No ponto de conexão com a Rede Básica que alimenta um só distribuidor ou consumidor, através de vários ramais, se não existir elementos de compensação reativa ou serviços auxiliares ligados no barramento, nem algum ramal capaz de alimentar consumidor livre, a medição pode ser global no lado de baixa do transformador.

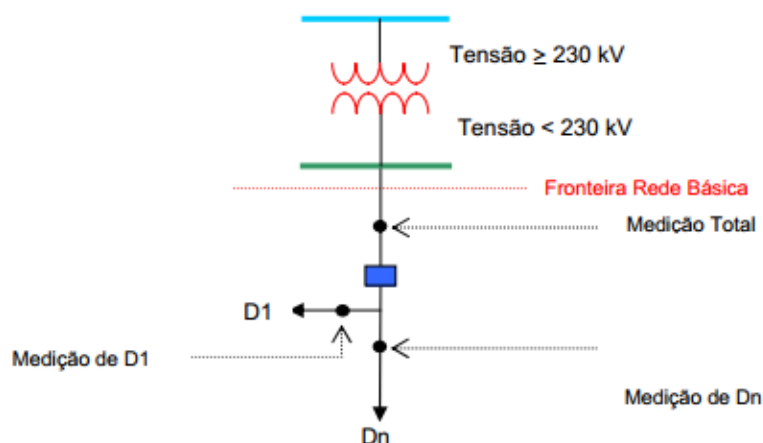
Figura 2.11 - Medição de Fronteira no Ponto de Conexão com a Rede Básica que alimenta um só distribuidor ou consumidor.



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

No ponto de conexão com a Rede Básica, cuja linha for compartilhada com mais de um distribuidor ou consumidor livre, a medição deve ser neste ponto e nos pontos de conexão de cada um.

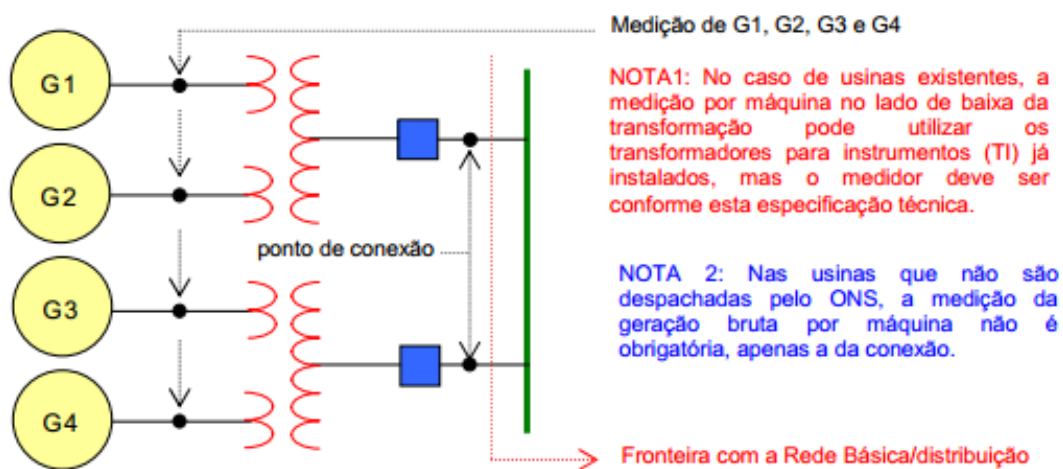
Figura 2.12 - Medição de Fronteira no Ponto de Conexão com a Rede Básica cuja linha é compartilhada por mais de um distribuidor ou consumidor livre



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

Nas usinas, cujas máquinas são agrupadas por transformador, a medição deve ser individual por máquina no lado de baixa do transformador e no ponto de conexão.

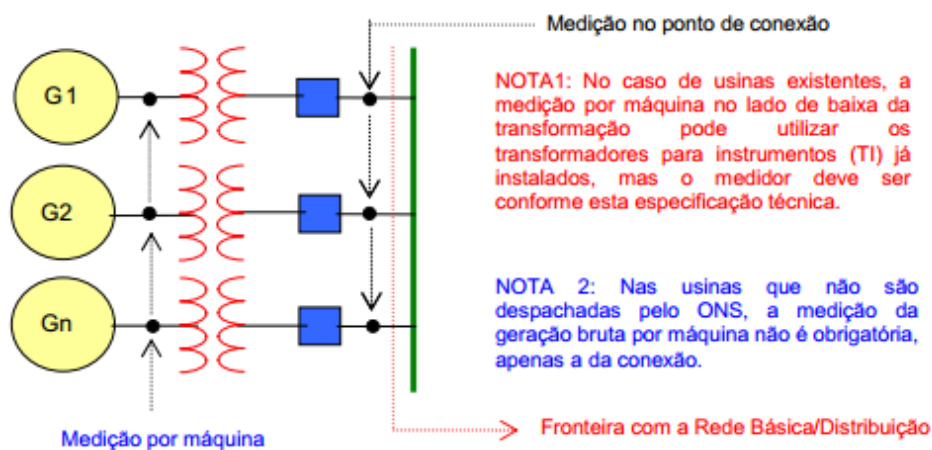
Figura 2.13 - Medição de Fronteira das Unidades Geradoras agrupadas por Transformador



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

Em usinas cujas máquinas são conectadas individualmente na Rede Básica ou na Rede de Distribuição, a medição deve ser por máquina no lado de baixa do transformador elevador e nos pontos de conexão.

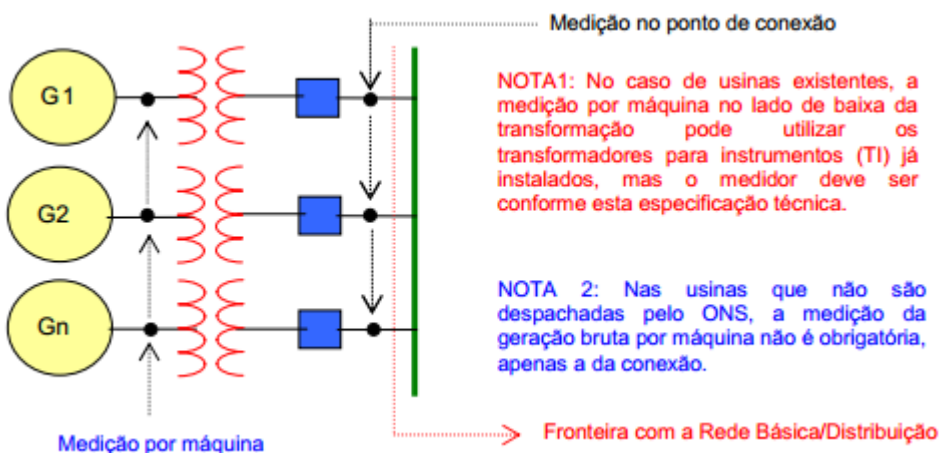
Figura 2.14 - Medição de Fronteira das Unidades Geradoras conectadas individualmente



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

Em usinas que se conectam à Rede Básica ou Rede de Distribuição através de linhas de transmissão de uso exclusivo, a medição deve ser por máquina no lado de baixa da transformação e nos pontos de conexão.

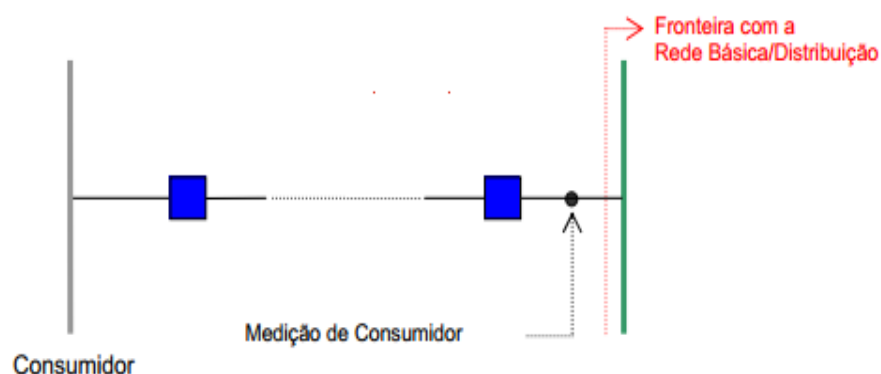
Figura 2.15 - Medição de Fronteira das Unidades Geradoras conectadas através de linhas de transmissão de uso exclusivo



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

Em consumidor, a medição deve ser no ponto de conexão com a Rede Básica ou Rede de Distribuição.

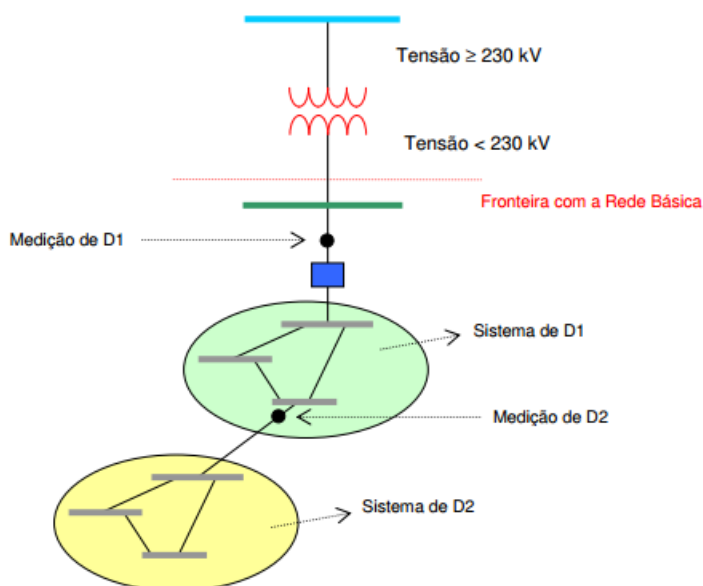
Figura 2.16 - Medição de Fronteira das Unidades Consumidoras



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

Nos pontos de conexão entre agentes que fazem parte da CCEE, a medição deve ser instalada no lado do agente conectado.

Figura 2.17 - Medição de Fronteira nos pontos de conexão de agentes distribuidores



Fonte – ONS_ Procedimento de Rede_Submódulo 12.2 (2004)

A etapa de leitura corresponde a localidade onde é feita a leitura e o dia em que ocorre este procedimento. Os dados do faturamento são muito importantes, pois é por intermédio desta informação, que podemos obter a quantidade de energia que não está sendo faturada naquele mês.

Figura 2.19 – Cadastro dos Transformadores de Distribuição e Perdas Correspondentes



Fonte – Enersul (2013)





















Com vistas à necessidade de obter informações complementares para direcionamento das ações de combate às perdas comerciais, foi implantada a medição para balanço energético em subestações, alimentadores e transformadores de distribuição (CP-Trafo).

A medição nos transformadores de distribuição é a forma mais restritiva de se apurar as perdas de uma região, já que o número de clientes atendidos por uma subestação ou alimentador é muitas vezes maior que o número de clientes atendidos pelos transformadores de distribuição.

O monitoramento dos transformadores é feito de forma simultânea com o seu alimentador correspondente. Inicialmente, verifica-se o alimentador com maior perda e, posteriormente, é investigado o correspondente transformador com maior percentual de perda.

As restrições são importantes no que se refere ao combate do furto de energia, já que o número de equipes de campo e o número de fiscalizações é bastante restrito. Ou seja, é inconcebível fiscalizar todas as unidades consumidoras no período de um mês, período este que serve como base para o cálculo das perdas comerciais.

Figura 2.20 - Cadastro do Transformador, Alimentador Correspondente e Perda do Trafo

	TRAFO	ALIMENTADOR	PERDA (kWh)	PORC %	MAPA
1	 401ET000295019 - CG505 -	Alimentador CG505	24.021,24	57,97%	Mapa
2	 401ET000057909 - CGA03 -	Alimentador CGA03	22.933,77	53,41%	Mapa
3	 401ET000216119 - CGA03 -	Alimentador CGA01	22.248,71	37,60%	Mapa
4	 401ET000046479 - CGL08 -	Alimentador CGL08	19.112,96	45,23%	Mapa
5	 401ET000246269 - CG503 -	Alimentador CG503	18.423,22	42,31%	Mapa
6	 401ET000065219 - CGM08 -	Alimentador CGM08	16.183,82	68,92%	Mapa
7	 401ET000046509 - CGL01 -	Alimentador CGL01	16.116,78	55,04%	Mapa
8	 401ET000063419 - CGA01 -	Alimentador CGA01	16.086,16	59,05%	Mapa
9	 401ET000045939 - CGL03 -	Alimentador CGL03	15.822,01	57,24%	Mapa
10	 401ET000046169 - CGA04 -	Alimentador CGA04	15.291,67	44,63%	Mapa
11	 401ET000061059 - CGA01 -	Alimentador CGA01	14.855,91	57,88%	Mapa
12	 401ET000057809 - CGA03 -	Alimentador CGA03	14.669,32	30,92%	Mapa
13	 401ET000039509 - CGA01 -	Alimentador CGA01	14.493,66	42,16%	Mapa
14	 401ET000092399 - CG505 -	Alimentador CG505	14.175,16	54,52%	Mapa
15	 401ET000057769 - CGA04 -	Alimentador CGA04	13.626,40	34,29%	Mapa
16	 401ET000207689 - CGL03 -	Alimentador CGL03	13.227,61	43,87%	Mapa
17	 401ET000243569 - CG503 -	Alimentador CG503	13.190,49	35,45%	Mapa
18	 401ET000046109 - CGL06 -	Alimentador CGL06	12.864,52	28,51%	Mapa
19	 401ET000042719 - CGL07 -	Alimentador CGL07	12.816,84	47,80%	Mapa
20	 401ET000084619 - CGA05 -	Alimentador CGA05	12.773,79	37,52%	Mapa

Fonte – Enersul (2013)

O código de comando da variável Perda Alimentador e Perda CP-Trafo é apresentado no quadro 2.5.

Quadro 2.5 - Linha de código plataforma Qlik View – Camada N1 – Perda Alimentador e Perda CP-Trafo

```
// Perda Alimentador
(
if ( num(PERDA_NT_%_AL_ano) >= 1, (1+2*(num(PERDA_NT_%_AL_ano))),1)
// Se a perda no alimentador for maior ou igual a 1, é atribuído o valor de (2*valor da perda) +1
//Caso contrário, o valor atribuído é 1
+
//Perda CP-Trafo
if ( num(PERDA_PERCENTUAL) >= 1, (1 + 3*(num(PERDA_PERCENTUAL))),1)
)
//Se a perda no CP-Trafo for maior ou igual a 1, é atribuído o valor de (3*valor da perda) +1
//Caso contrário, o valor atribuído é 1
```

Fonte – Elaborado pelo autor (2014)

É seguro afirmar que o valor de perda no CP-Trafo tem um peso maior, comparado ao valor da perda no alimentador (2), pois a informação de perda no transformador é mais restritiva e confiável. A perda no alimentador abrange uma área grande de UC's fraudadoras, dificultando a assertividade das fiscalizações. Já o transformador abrange uma área menor de clientes, tornando mais fácil a busca pelo suspeito. Sendo assim, o valor que a variável CP-Trafo tem na composição da camada N1 é de grande importância.

2.4.7 Camada N2

A camada **N2** é composta por quatro variáveis: Pirâmide, Variação de Consumo, Troca de Titularidade e Inspeções Recentes. Nesta camada as variáveis são multiplicadas, o que demonstra maior correlação entre elas. Se uma variável sofre qualquer tipo de variação, o resultado reflete nas demais.

Quadro 2.6 – Composição da Camada N2

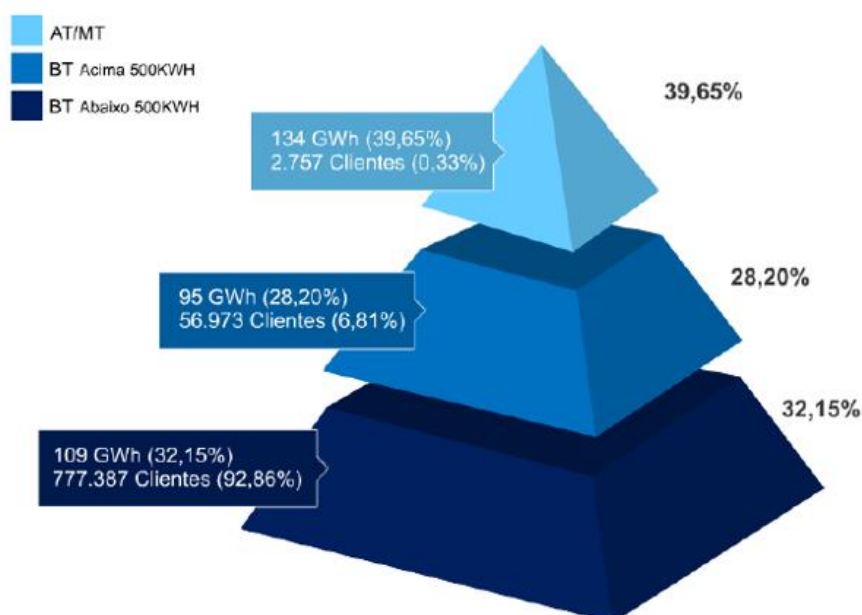
//Pirâmide
*
//Variação de Consumo
*
//Troca de Titularidade
*
//Inspeções Recentes

Fonte: Elaborado pelo autor (2014)

2.4.8 Pirâmide – Estratégia por Segmentação

Objetivando uma ação focada e eficaz das fiscalizações, surgiu o planejamento estratégico por segmentação de acordo com a energia consumida, direcionando medidas específicas por segmento de mercado, sempre priorizando a quantidade de energia agregada por números de fiscalizações.

As prioridades das fiscalizações são dadas aos maiores mercados da Enersul, ou seja aqueles que após serem regularizados darão maior retorno de energia agregada e a recuperação de receita cobrirá pelo menos os custos com fiscalizações.

Figura 2.21 - Pirâmide do Modelo de Segmentação de Fiscalizações

Fonte – Enersul (2014)

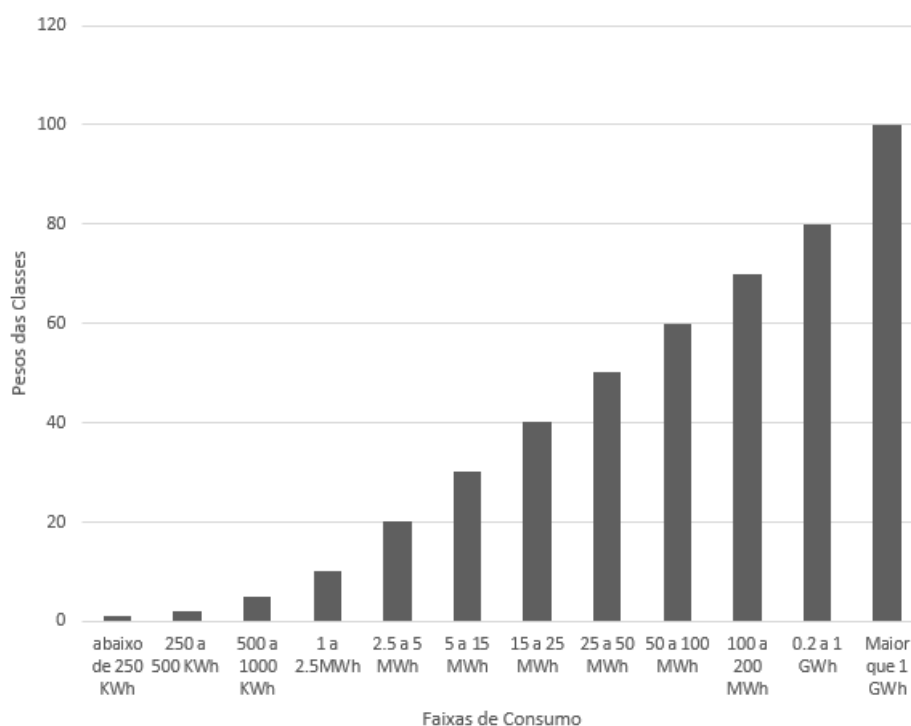
O código de comando da variável Pirâmide é apresentado no quadro 2.7. O gráfico 2.2 demonstra o nível de importância de todas as faixas de consumo.

Quadro 2.7 - Linha de código plataforma Qlik View – Camada N2 - Pirâmide

```
// Perda Alimentador
if (Faixa='Maior 1 GWh',100, // Se a faixa de consumo for maior que 1 GWh, atribui-se peso 100, senão
if (Faixa = '0.2 a 1 GWh',80, // Se a faixa de consumo estiver entre 0.2 e 1 GWh, atribui-se peso 80, senão
if (Faixa = '100 a 200 MWh',70, // Se a faixa de consumo estiver entre 100 e 200 MWh atribui-se 70, senão
if (Faixa = ' 50 a 100 MWh',60, // Se a faixa de consumo estiver entre 50 e 100 MWh atribui-se 60, senão
if (Faixa = ' 25 a 50 MWh',50, // Se a faixa de consumo estiver entre 25 e 50 MWh atribui-se 50, senão
if (Faixa = ' 15 a 25 MWh',40, // Se a faixa de consumo estiver entre 15 e 250 MWh atribui-se 40, senão
if (Faixa = ' 5 a 15 MWh',30, // Se a faixa de consumo estiver entre 5 e 15 MWh atribui-se 30, senão
if (Faixa = ' 2.5 a 5 MWh',20, // Se a faixa de consumo estiver entre 2.5 e 5 MWh atribui-se 20, senão
if (Faixa = ' 1 a 2.5 MWh',10, // Se a faixa de consumo estiver entre 1 e 2.5 MWh atribui-se 10, senão
if (Faixa = ' 500 a 1000 KWh',5, // Se a faixa de consumo estiver entre 500 e 1000 KWh atribui-se 5,, senão
if (Faixa = ' 250 a 500 KWh',2, // Se a faixa de consumo estiver entre 250 e 500 KWh atribui-se 2, senão
if (Faixa = ' abaixo de 250 KWh',1, // Se a faixa de consumo estiver abaixo de 250 KWh atribui-se 1, senão
1)))))))))) // Se nenhuma das hipóteses anteriores é satisfeita, atribui-se peso 1
```

Fonte: Elaborado pelo autor (2014)

Gráfico 2.2 - Pesos atribuídos para cada faixa de consumo



Fonte: Elaborado pelo autor (2014)

2.4.9 Variação de Consumo

A variação de consumo é um fator determinante para a composição da segunda camada. Muitas unidades consumidoras fraudadoras apresentam um comportamento de consumo característico após a irregularidade.

Após ser feita a fraude pode ocorrer uma queda acentuada no consumo, o que configura a irregularidade. Fatores climáticos, épocas do ano e período de férias podem causar uma queda no consumo no histórico do cliente, sendo assim deve-se ter o cuidado de não atribuir fiscalizações para essas situações.

Quadro 2.8 - Linha de código plataforma Qlik View – Camada N2 – Variação de Consumo

```
// Variação de Consumo

(1+
if(
if ( num$(VARIACAO_CONSUMO)) <= '0.7'           //Se a variação de consumo for menor ou igual a 0.7
and                                              // e
Ano_Sit <= '2012'                               // O ano da ligação da UC ocorreu antes de 2012
and                                              // e
CodSitUC = 'LG'                                 // A unidade consumidora encontra-se ligada
And                                              // e
num (Var_ClasseConsumo) > '1',                 //A variação em relação a sua classe de consumo for
maior que 1
(1/num$(VARIACAO_CONSUMO)))                    // Então adota-se o valor (1/ valor da variação de
consumo)
*                                              // O resultado da operação anterior multiplicado por
(1+0.01*$(CONSUMO_MAXIMO)),                    // (1+ um por cento do valor do consumo máximo da unidade)
0)                                              // Se uma das condições não forem aceitas, adota-se o valor 0
*                                              // O resultado da operação anterior multiplicado por
if (num$(VARIACAO_CONSUMO))> 1, 0.01,1)//0.01 se o valor da variação de consumo for maior que 1
                                              // Se a variação não for maior que 1, o valor adotado será 1
*                                              // O resultado anterior multiplicado por
if (Ano_Sit > 2012, 0.01, 1)                   // 0.01 se a ligação da UC for maior que 2012, senão o valor é 1
```

Fonte – Elaborado pelo Autor (2014)

2.4.10 Troca de Titularidade

Novos consumidores devem ser investigados, pois existe a probabilidade do novo morador da unidade consumidora fazer a fraude já no primeiro mês do seu faturamento. A irregularidade provocada pelo novo morador é difícil de ser detectada, já que o novo cliente pode apresentar hábitos de consumo diferentes comparados aos hábitos do cliente anterior. Portanto a queda de consumo da nova unidade consumidora pode retratar um hábito normal do novo consumidor, não configurando uma irregularidade.

Uma estratégia de investigação para trocas de titularidade é investigar as unidades que tiveram trocas antigas, no caso anteriores a 2013, já que titulares antigos e sem nenhum histórico de fiscalização podem ser potenciais fraudadores.

Quadro 2.9 - Linha de código plataforma Qlik View – Camada N2 – Troca de Titularidade

```
// Troca de Titularidade
if ( ANO_TT > `2013`, 0.01, 1) //Se o ano em que ocorreu a última troca de titularidade for maior que 2013, o peso atribuído será 0.01, se o ano da troca for menor ou igual a 2013, atribui-se o peso 1.
```

Fonte : Elaborado pelo Autor (2014)

2.4.11 Inspeções Recentes

Unidades consumidoras com inspeções recentes são descartadas da hipótese de fraude, sendo assim a prioridade para realização de fiscalizações é dada para unidades que tenham sido fiscalizadas há mais de um ano.

Se as inspeções recentes forem realizadas há menos de um ano ou a última inspeção recente foi considerada normal (código 005), então o peso atribuído para esta variável é bem pequeno. Caso a unidade tenha tido inspeções há mais de um ano e tenha tido um parecer irregular, o valor do peso atribuído é 1.

Quadro 2.10 - Linha de código plataforma Qlik View – Camada N2 – Inspeções Recentes

```
// Inspeções Recentes
if (( max(ANO_FIS) >2013 or max(Ano_Sem_Atendimento) > 2013 or //Se a última fiscalização ocorreu no ano de 2014 ou o ano sem atendimento for maior que 2013 ou
(max(ANO_FIS) >= 2013 and //Se a última fiscalização ocorreu em 2013 ou 2014
(only (if ( Data_Formatada = MAX_DATA, [Cod Parecer])) = `005`)), 0.0001,1) // E somente se o código do parecer for normal, atribui-se o valor 0.0001, caso contrário atribui-se o valor 1.
```

Fonte: Elaborado pelo Autor (2014)

2.4.12 Camada N3

A camada N3 é formada por apenas duas variáveis, são elas: Tipo de Fase e Classe Principal. O fato das duas variáveis estarem multiplicadas demonstra a importância que as duas juntas têm na composição da camada N3.

Quadro 2.11 – Composição da Camada N3

```
// Tipo de Fase
*
// Classe Principal
```

Fonte: Elaborado pelo Autor (2014)

2.4.13 Classe Principal

Quadro 2.12 - Linha de código plataforma Qlik View – Camada N3 – Classe Principal

```
// Classe Principal
if(COD_CLAS_PRIN = '02', 5, //Se o código da classe for 02 (industrial) é atribuído o peso 5
if(COD_CLAS_PRIN = '03', 4, //Se o código da classe for 03 (comercial) é atribuído o peso 4
if(COD_CLAS_PRIN = '04', 3, //Se o código da classe for 04 (rural) é atribuído o peso 3
if(COD_CLAS_PRIN = '01', 2, //Se o código da classe for 01 (residencial) é atribuído o peso 2
if (COD_CLAS_PRIN = '05' or COD_CLAS_PRIN = '07', 1, //Se o código da classe for 05 (poder público) ou o código da classe for 07 (serviço público) é atribuído o peso 1
0)))) //Se alguma UC não estiver enquadrada em nenhuma das classes o peso atribuído será zero.
```

Fonte: Elaborado pelo Autor (2014)

As classes consumidoras são divididas em sete, de forma que todos os consumidores atendidos pela concessionária devem ser enquadrados em um tipo de classe principal. O peso atribuído à esta variável depende do tipo da classe. O código de maior peso é atribuído à classe Industrial (código 02). Por meio da experiência do operador e dos resultados das fiscalizações constatou-se que as autuações que geraram maior retorno foram realizadas em consumidores industriais, os quais desviam maior quantidade de energia.

O algoritmo referente à classe principal define que o maior peso é atribuído para a classe 02, industrial, cujo valor é 5. A segunda maior classe e mais importante é a classe comercial 03, com peso 4. Os consumidores rurais, classe 04, são os terceiros mais importantes e recebem peso 3. A quarta classe e não menos importante é a classe residencial, a maior de toda a concessão Enersul e recebe o peso 2. Por fim, as classes 05 e 07 referentes a poder e serviço público recebem peso 1. Caso exista algum consumidor que não se enquadre em algumas das classes, deve-se atribuir a essa unidade consumidora o peso 0.

As divisões dos consumidores são feitas por intermédio de classes principais.

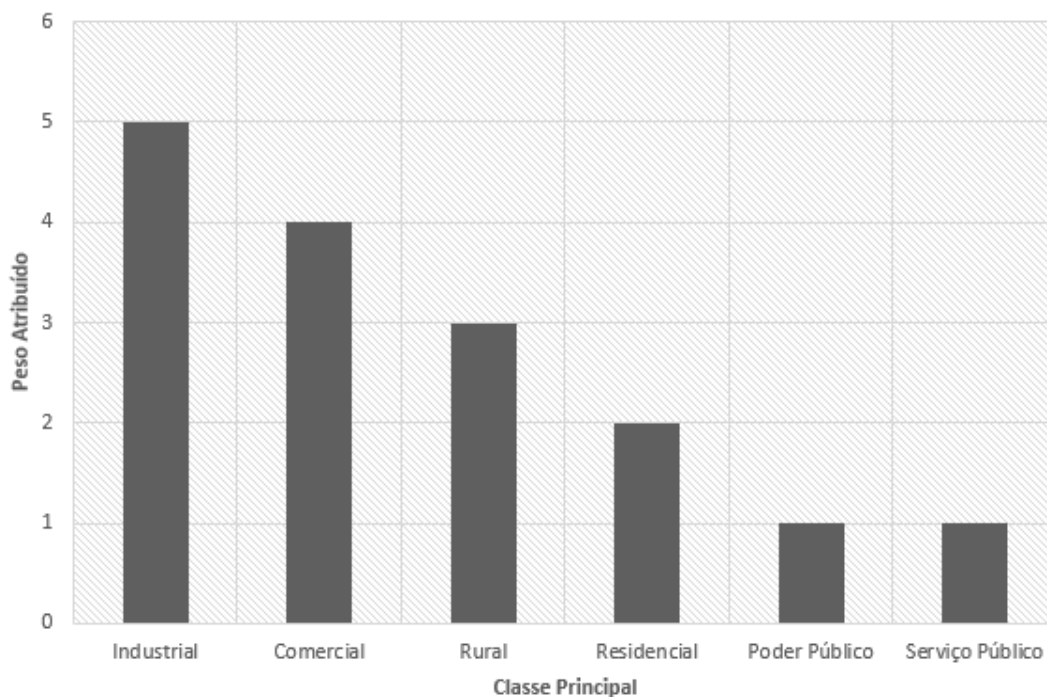
Tabela 2.2- Código da Classe Principal Consumidora

Classe Principal	Código da Classe	Peso
Industrial	02	5
Comercial	03	4
Rural	04	3
Residencial	01	2
Poder Público	05	1
Serviço Público	07	1

Fonte: Elaborado pelo autor (2014)

A classe consumidora serviço público refere-se a todos os serviços prestados à sociedade, que são da iniciativa privada, como por exemplo abastecimento de água e esgoto. A classe poder público refere-se a todos os órgãos da esfera pública e que são administrados pelo Estado.

O gráfico 2.3 mostra o nível de contribuição de cada classe de consumo, na constituição da camada N3.

Gráfico 2.3 - Pesos atribuídos para cada faixa de consumo

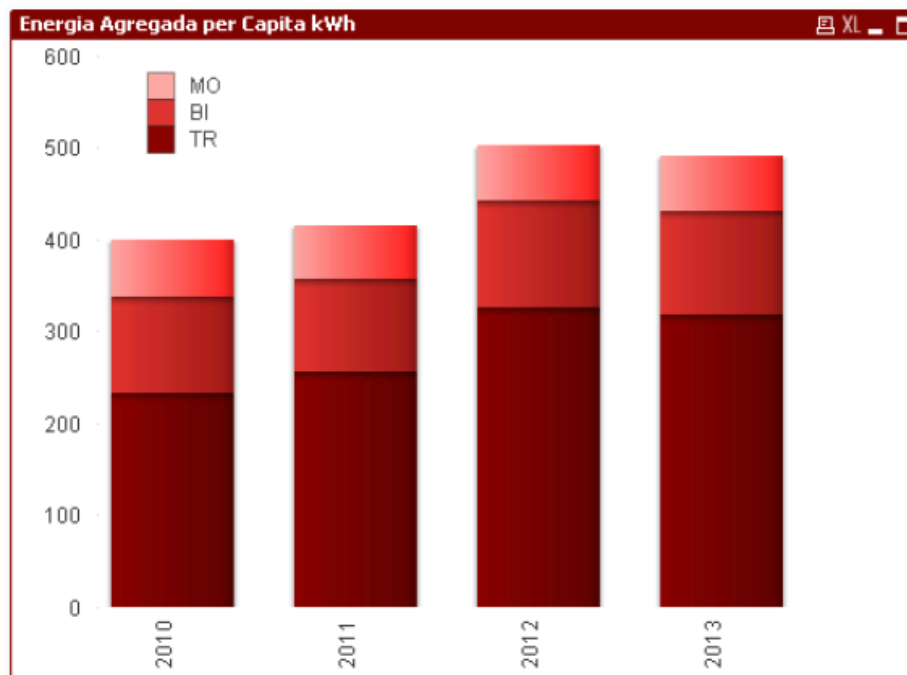
Fonte: Elaborado pelo autor (2014)

2.4.14 Tipo de Fase

Outra variável importante a respeito dos consumidores que compõem a terceira camada é o tipo de fase de alimentação, ou seja clientes monofásicos, bifásicos ou trifásicos. Os consumidores trifásicos são os que contribuem com a maior parcela da energia agregada e também apresentam alto índice de reincidência na prática das irregularidades. Atualmente, a maioria das unidades consumidoras atendidas pela Enersul é do grupo de tensão monofásico, somando um total de 735.528 consumidores. O segundo maior grupo pertence à classe bifásica, com um total de 397.852 UC's. O menor grupo é formado pelos consumidores trifásicos, 199.083 clientes.

Mesmo com um número menor de consumidores trifásicos, estes clientes consomem grande parte da energia distribuída e são os que contribuem com a maior parcela da energia agregada. Portanto, autuações feitas nesta faixa de atendimento são muito importantes e têm maior peso na composição do DNA.

Gráfico 2.4 – Energia Agregada Per Capita



Fonte: Enersul (2014)

O gráfico 2.4 retrata a evolução do montante da energia agregada referente aos anos de 2010 a 2013. É possível constatar que houve um progresso no que se refere a quantidade de energia agregada ao sistema, graças a implantação da metodologia que utiliza a detecção das irregularidades com auxílio do algoritmo do DNA da fraude.

A energia agregada per capita evoluiu 22% (comparação do biênio 2013 – 2012 versus 2011 – 2010) em clientes do Grupo B, consequência direta do alinhamento estratégico da autuação dos clientes de maior carga, no caso trifásicos. A quantidade de fiscalizações evolui apenas 4% (comparação do biênio 2013 – 2012 versus 2011 – 2010), demonstrando que os investimentos em tecnologia para aumentar a assertividade nas inspeções surtiram efeito.

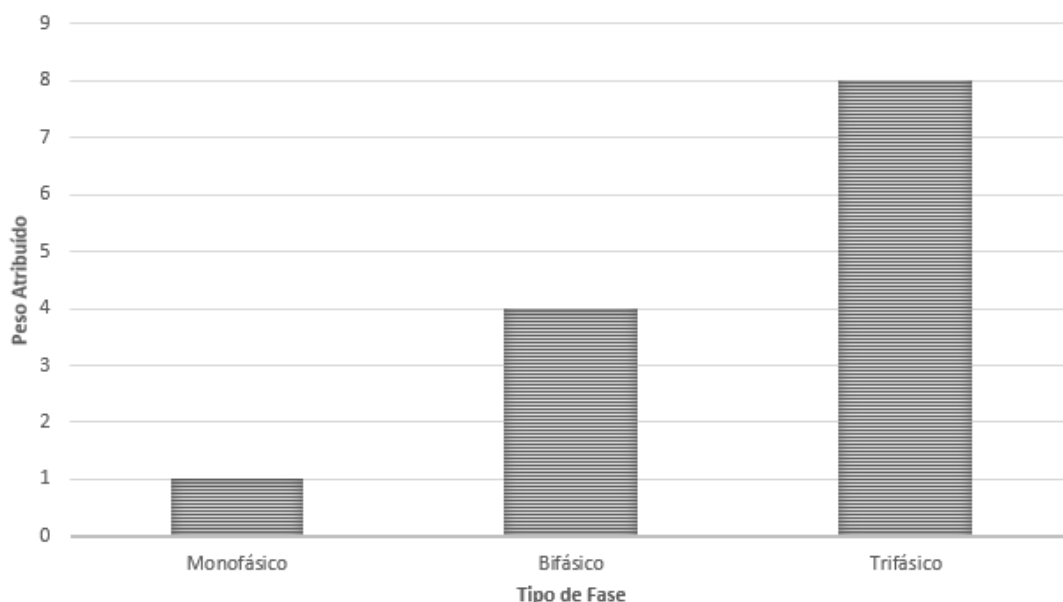
No algoritmo do DNA, os clientes trifásicos recebem um peso mais significativo, no caso 8. Para os clientes bifásicos é atribuído o peso 4 e para clientes monofásicos o peso é 1. O maior peso determina a importância do tipo de fase para o valor de energia agregada.

Quadro 2.13 - Linha de código plataforma Qlik View – Camada N3 – Tipo de Fase

```
// Tipo de Fase
(1+ // Soma-se 1 para evitar que o resultado das operações dê zero.
if (GrupoFat = 'B' and CodFase = 'TR', 8, 0)) // Se o grupo de faturamento for B e o código da fase for
trifásico, então o valor 8 é atribuído, senão o valor é zero.
* // O resultado anterior vezes
(1+ // Soma-se 1 para evitar que o resultado das operações dê zero.
if (GrupoFat = 'B' and CodFase = 'BI', 4, 0)) // Se o grupo de faturamento for B e o código da fase for
bifásico, então o valor 4 é atribuído, senão o valor é zero
* // O resultado anterior vezes
(1+ // Soma-se 1 para evitar que o resultado das operações dê zero.
if (GrupoFat = 'B' and CodFase = 'MO', 1, 0)) // Se o grupo de faturamento for B e o código da fase for
monofásico, então o valor 1 é atribuído, senão o valor é zero.
```

Fonte: Elaborado pelo autor (2014)

O gráfico 2.5 detalha o nível de importância das três classes consumidoras para a formação da terceira camada. Os consumidores trifásicos têm um nível de importância oito vezes maior que os clientes monofásicos.

Gráfico 2.5 – Peso atribuído a cada tipo de fase

Fonte: Elaborado pelo autor (2014)

2.4.15 Camada N4

A camada N4 é formada pelas variáveis reincidência e energia agregada.

Quadro 2.14 – Composição da Camada N4

// Energia Agregada
*
// Reincidentes

Fonte: Elaborado pelo autor (2014)

2.4.16 Energia Agregada

O termo energia agregada refere-se ao montante de energia, que torna a ser faturada após uma fiscalização, ou seja a energia que não estava sendo computada, devido à irregularidade.

O peso desta variável dependerá da quantidade de energia agregada, que foi inserida no sistema, referente à última fiscalização da unidade consumidora. Quanto menor for o valor da energia agregada, maior será o peso desta variável. Isso porque, após a regularização da UC, a tendência é que se tenha alto valor de energia agregada. Como o valor desta energia, após a última fiscalização é pequeno a fraude pode ter sido refeita, não havendo aumento do consumo após a regularização.

Quadro 2.15 - Linha de código plataforma Qlik View – Camada N4 – Energia Agregada

```
// Energia Agregada
if ($ (Energia_Agregada_MWH) < 0, 2, //Se a energia agregada for menor que zero, atribui-se peso 2
if ($ (Energia_Agregada_MWH) > 0, 0.01, //Se a energia agregada for maior que zero, atribui-se peso 0.01
if ($ (Energia_Agregada_MWH) = 0, 1.1.))) //Se a energia agregada for igual a zero, atribui-se peso 1.1
```

Fonte: Elaborado pelo autor (2014)

2.4.17 Reincidentes

Quadro 2.16 - Linha de código plataforma Qlik View – Camada N4 – Reincidentes

```

// Reincidentes
if (UC_Autuada, (1+1*NumericCount (UC_Autuada)),1) //Se a unidade consumidora tiver sido autuada, o valor atribuído será 1+1*(o número de autuações), se não houve autuações o valor atribuído será 1.
* // O resultado anterior vezes
if ((max(ANO_FIS)>=2013 or max(Ano_Sem_Atendimento)>=2013) and //Se a última fiscalização aconteceu entre 2013 e 2014 ou o ano de atendimento ocorreu antes de 2013 e
(only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='005'),0.01, // Somente se o parecer da última fiscalização for 005 (normal), o valor obtido anteriormente será multiplicado por 0.01, senão
if ((max(ANO_FIS)>=2013 and //Se a última fiscalização ocorreu entre 2013 e 2014 e
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='001' or // Somente se o parecer da última fiscalização for 001 (porta/portão fechado), senão
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='002' or // Somente se o parecer da última fiscalização for 002 (cão bravo solto),senão
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='003' or // Somente se o parecer da última fiscalização for 003 (medidor ilegível),senão
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='010'), 3, // Somente se o parecer da última fiscalização for 010 (lacre violado), para quaisquer destes códigos atribui-se o peso 3.
if ((max(ANO_FIS)>=2012 and max(ANO_FISC)<2013 and //Se a última fiscalização for em 2012 e
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='005', 0.001, // Somente se o parecer tiver sido 005 (normal), o peso atribuído será 0.001, senão
if ((max(ANO_FIS)>=2011 and max(ANO_FISC)<2012 and //Se a última fiscalização for em 2011 e
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='005', 0.002, // Somente se o parecer tiver sido 005 (normal), o peso atribuído será 0.002, senão
if ((max(ANO_FIS)>=2010 and max(ANO_FISC)<2011 and //Se a última fiscalização for em 2010 e
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='005', 0.004, // Somente se o parecer tiver sido 005 (normal), o peso atribuído será 0.004, senão
if ((max(ANO_FIS)>=2009 and max(ANO_FISC)<2010 and //Se a última fiscalização for em 2009 e
only (if (Data_Formatada = MAX_DATA, [Cod_Parecer]))='005', 0.005, // Somente se o parecer tiver sido 005 (normal), o peso atribuído será 0.005
1)))))) //Se nenhuma das restrições anteriores forem aceitas, o valor atribuído será 1.

```

Fonte: Elaborado pelo autor (2014)

A reincidência é uma característica típica de fraudadores, sendo que unidades consumidoras que já tiveram fiscalizações anteriores têm grandes chances de estar novamente fraudando. Muitas fiscalizações demonstram que os clientes autuados, já tinham sido pegos anteriormente.

A variável Reincidentes leva em consideração não só o número de autuações da unidade consumidora, mas também o quão distante foram estas autuações. Um fato interessante é que algumas unidades consumidoras, que tiveram parecer normal no passado são autuadas com fraude. Sendo assim, mesmo que o parecer seja normal para autuações passadas, quanto mais antigo for este parecer, maior peso haverá na composição da variável reincidência.

2.4.18 Camada N5

A camada N5 é formada pelo histórico de contas do cliente, sendo que bons pagadores tem menor relevância no peso da quinta camada. Essa camada é formada pelos fatores, contas vencidas, débito acumulado, classe principal e inadimplência.

Quadro 2.17 – Composição da Camada N5

```

if (Count(UC_Autuada) > 0 and // Se a UC tiver sido autuada pelo menos uma vez e
CodSitUC = 'DS' and          // Estiver desligada no Sistema e
Ano_Sit >= 2009,             // As autuações ou o desligamento aconteceram entre
                             // 2009 e 2014, então os cálculos seguintes são realizados
                             // Fatura Média
                             +
                             // Contas Vencidas
                             *
                             // Reincidência
                             +
                             // Débito Acumulado
                             *
                             // Classe Principal
                             +
                             // Inadimplência

```

Fonte: Elaborado pelo autor (2014)

É válido destacar que as variáveis contas vencidas e reincidência tem seus resultados multiplicados, demonstrando a forte ligação entre elas. O mesmo acontece com as variáveis débito acumulado e classe principal. Os maiores clientes, comerciais e industriais, recebem os maiores pesos, no entanto o tipo de classe só será relevante se houver grandes quantias de débitos acumulados, mostrando a correlação entre estas duas variáveis.

2.4.19 Fatura Média

O termo fatura média refere-se ao valor total de todas as faturas que não foram quitadas dividido pelo número destas faturas. Dez por cento desse valor é o peso atribuído a esta variável.

Quadro 2.18 – Linha de código plataforma Qlik View – Camada N5 – Fatura Média

```
// Fatura Média
0.1*(sum(Valor_Fat) / Count (UC_Inadimplencia) // Divide-se o valor total dos débitos acumulados da unidade pelo número de faturas inadimplentes. Dez por cento deste resultado é o valor da variável fatura média.
```

Fonte: Elaborado pelo autor (2014)

2.4.20 Contas Vencidas

A soma do número de contas vencidas da unidade consumidora determinará o peso desta variável.

As unidades consumidoras com débitos são cortadas em campo e permanecem desligadas no sistema até futuro pagamento dos débitos e posterior religação. Muitos clientes cortados efetuam a ligação clandestina, denominada auto religação. Sendo assim é preciso rastrear as UC's que têm contas vencidas, pois é bem provável que dentre elas possa existir ligações clandestinas

Quadro 2.19 – Linha de código plataforma Qlik View – Camada N5 – Contas Vencidas

// Contas Vencidas Count(UC_Inadimplencia)	<i>// O número total de contas vencidas da UC</i>
--	---

Fonte: Elaborado pelo autor (2014)

2.4.21 Débito Acumulado

A variável débito acumulado tem um peso pequeno, já que muitos clientes nesta situação apresentam pequeno histórico de fraude. Muitos deles procuram a concessionária para refinarçar suas dívidas ou quitar seus dividendos. No entanto, a ocorrência de irregularidades em clientes que apresentam débito acumulado continua acontecendo, mas em escala menor se comparada, por exemplo com clientes reincidentes.

Quadro 2.20 – Linha de código plataforma Qlik View – Camada N5 – Débito Acumulado

// Débito Acumulado 0.001*sum(Valor_Fat)	<i>// 0.1% do valor total do débito acumulado</i>
--	---

Fonte: Elaborado pelo autor (2014)

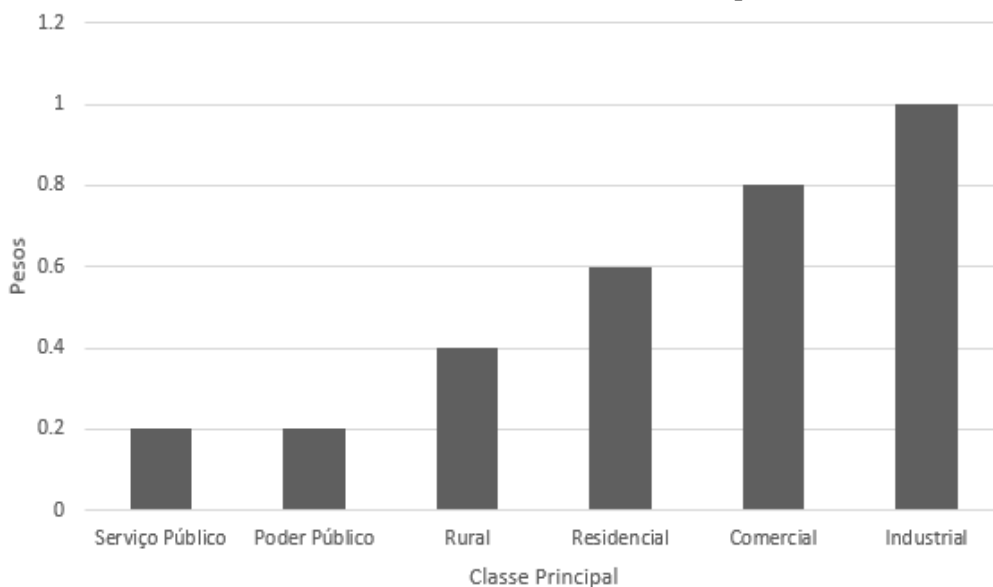
2.4.22 Classe Principal

A camada N5 leva em consideração a variável classe principal destacando a importância dos consumidores industriais, os quais tem um peso máximo atribuído, 1. Em ordem de importância temos os consumidores industriais, seguidos dos comerciais, residenciais, rurais e por último poder e serviço público.

Quadro 2.21 – Linha de código plataforma Qlik View – Camada N5 – Classe Principal

// Classe Principal if (COD_CLAS_PRIN = '02', 1.0, <i>// Se a classe principal for industrial, o peso é 1</i> if (COD_CLAS_PRIN = '03', 0.8, <i>// Se a classe principal for comercial, o peso é 0.8</i> if (COD_CLAS_PRIN = '01', 0.6, <i>// Se a classe principal for residencial, o peso é 0.6</i> if (COD_CLAS_PRIN = '04', 0.4, <i>// Se a classe principal for rural, o peso é 0.4</i> if (COD_CLAS_PRIN = '05' or <i>// Se a classe principal for poder público ou</i> COD_CLAS_PRIN = '07', 0.2, <i>// Se a classe principal for serviço público, o peso é 0.2</i> 0)))) <i>// Se a unidade consumidora não estiver enquadrada em nenhuma das classes atribui-se o peso 0.</i>
--

Fonte: Elaborado pelo autor (2014)

Gráfico 2.6 – Peso atribuído às Classes Principais

Fonte: Elaborado pelo autor (2014)

2.4.23 Inadimplência

O fator inadimplência é um grande agravante no que se refere às irregularidades. Um número expressivo de consumidores, que têm seu fornecimento de energia cortado recorre aos furtos de energia.

A perda de poder aquisitivo da população brasileira e a diminuição da capacidade de arcar com compromissos básicos, tem se acentuado no Brasil, nos últimos 10 anos. A Associação Brasileira de Distribuidores de Energia Elétrica, ABRADEE, mostrou um crescimento da quantidade de faturas atrasadas superior a dez dias [10]. O atraso no pagamento de faturas não é considerado como inadimplência, no entanto este dado reflete a redução do poder aquisitivo dos clientes. Os constantes aumentos nas tarifas de energia elétrica, altos níveis de desemprego e a diminuição da renda são fatores preponderantes não só para o aumento da inadimplência, como também para a recorrência a atos fraudulentos.

Devido à importância desta variável, o valor do seu peso na composição do DNA será igual à soma dos valores de todas as faturas pendentes da unidade consumidora dividido por 1000.

Quadro 2.22 – Linha de código plataforma Qlik View – Camada N5 – Inadimplência

```
// Inadimplência
sum ( Valor_Fat)/ 1000      //Soma de todas as faturas sem pagamento, dividido por 1000.
```

Fonte: Elaborado pelo autor (2014)

2.5 Estrutura do Algoritmo DNA

Após a decodificação de todas as variáveis e da compreensão do significado de cada uma delas é importante destacar o processo de obtenção dos pesos de cada camada e de cada variável. Como não foi possível obter maiores informações a respeito de como o algoritmo foi criado, algumas conclusões sobre sua criação foram tomadas por meio da análise da estrutura do código. É possível concluir que cada variável, que está presente nas cinco camadas tiveram os pesos atribuídos por meio da experiência do operador e do feedback das equipes de campo.

Após agrupar as variáveis em camadas, um peso é adotado para cada camada, que melhor representa a importância desta camada na formação do score. Por meio de uma regressão linear simples é possível encontrar os pesos mais adequados para realização do somatório, já que o modelo do algoritmo é simples e não possui funções de ativação.

As funções de ativação geram uma saída a partir dos dados de entrada e de seus respectivos pesos. Os tipos destas funções podem ser degrau, sigmoideal, linear ou gaussiana [11]. Como nosso algoritmo não contém funções de ativação, é possível que os pesos das camadas tenham sido encontrados por meio de regressão linear simples.

A análise de regressão linear simples tem o propósito de avaliar o impacto relativo de uma variável sobre um resultado particular [12]. O estudo em questão pretende avaliar a influência de cada camada na determinação do perfil fraudador. Como não existem registros de como os pesos das variáveis e das camadas foram obtidos, uma análise hipotética será feita para demonstrar o método possivelmente empregado. É bem provável que um estudo simples de regressão linear deva ter sido utilizado, devido sua simplicidade e eficiência.

2.5.1 Determinação dos parâmetros do sistema.

Uma amostra com 87 unidades consumidoras inspecionadas e com o parecer conhecido foi analisada. Deseja-se determinar o valor dos pesos de cada variável e a importância que cada um deles exerce na determinação do perfil fraudador.

Como não há registros nem informações de como foram encontrados os pesos atribuídos às camadas, uma análise hipotética será feita com o objetivo de simular a técnica de regressão linear simples que pode ter sido utilizada na determinação dos pesos das camadas.

As variáveis utilizadas para análise são denominadas preditoras. Neste estudo 7 variáveis foram disponibilizadas pela empresa Enersul, para o estudo de caso. São elas:

- Consumo Máximo: Equivale ao máximo valor de energia ativa consumida da UC ao longo dos doze últimos ciclos de faturamento;
- Ano Situação: Corresponde ao ano que a unidade consumidora foi ligada;
- Classe Residencial: Classifica todas as unidades consumidoras da classe residencial;
- Classe Comercial: Classifica todas as unidades consumidoras da classe comercial;
- Peso dos Apontamentos: Determina os valores dos pesos atribuídos às unidades que tiveram apontamento do leiturista;
- Média dos Doze Faturamentos: Corresponde à média dos doze últimos ciclos completos de faturamento.
- Score: Valor resultante obtido do algoritmo do DNA da Fraude.

A primeira análise a ser feita é a determinação dos pesos de cada preditora na determinação do perfil fraudador.

Toda a análise e os cálculos serão feitos no software Matlab®, que calcula o valor exato dos pesos de cada preditora e determina qual delas influencia mais na determinação do resultado esperado.

2.5.2 Normalização dos dados

Num primeiro momento, os dados de cada preditora são os mais diversos, formados por números, textos e datas. Para obtenção de resultados mais fidedignos é interessante colocar todos os dados numa mesma base e referenciá-los numa mesma escala.

A tabela 2.3 mostra os valores das preditoras de duas unidades consumidoras. Uma delas com fraude e a outra normal.

Tabela 2.3 – Preditoras e Resultado do Parecer

Consumo Máximo	Ano Situacao	Residencial	Comercial	Peso Apontamento	Média dos 12 faturamentos	Score	PARECER
422	2007	Sim	Não	6	42	245	Fraude
807	2012	Não	Sim	1	10	235	Normal

Fonte: Elaborado pelo autor (2014)

Os dados das preditoras aparecem nas mais variadas formas. Dessa maneira é preciso transformá-los em números e posteriormente normalizá-los.

Tabela 2.4 – Dados Convertidos em Números

Consumo Máximo	Ano Situacao	Residencial	Comercial	Peso Apontamento	Média dos 12 faturamentos	Score	PARECER
422	2007	1	0	6	42	245	201
807	2012	0	1	1	10	235	0

Fonte: Elaborado pelo autor (2014)

Os termos sim e não foram considerados como valores 1 e 0, respectivamente. A UC com fraude recebe o valor 201 (score mínimo para fraude) e o valor 0 foi atribuído para não fraudador.

No Matlab®, devemos criar matrizes colunas de todas as preditoras, já representadas por números.

Figura 2.22- Matrizes coluna das preditoras

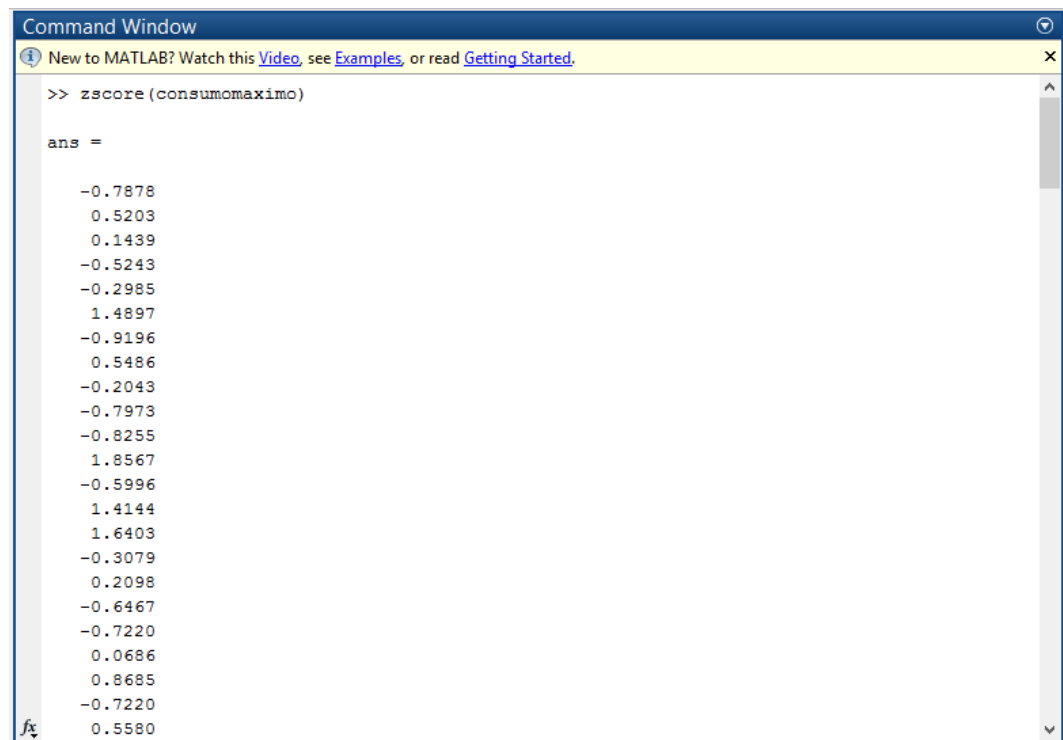
anosituacao	<87x1 double>
comercial	<87x1 double>
consumomaximo	<87x1 double>
media12fat	<87x1 double>
pesoapontamento	<87x1 double>
residencial	<87x1 double>
score	<87x1 double>

Fonte: Matlab® (2014)

Após declarar todas as variáveis é preciso normalizar os dados, para obtenção de valores de pesos de mesma escala.

O comando utilizado no Matlab® para normalização dos dados é **zscore**. Portanto, aplicando o comando zscore em todas as matrizes, encontra-se as matrizes normalizadas de todas as preditoras, inclusive a matriz normalizada do parecer.

Figura 2.23 - Exemplo de utilização do comando zscore



Fonte: Matlab® (2014)

2.5.3 Determinação dos Pesos das Predictoras

Após a normalização das predictoras e do parecer desejado, os pesos de cada variável preditora será determinado. Estes pesos refletem a importância de cada variável sobre o resultado desejado, no caso o parecer.

Para determinação dos pesos é utilizado o comando **regress**, que calcula o peso ideal de cada preditora para a obtenção da saída desejada. As variáveis normalizadas foram intituladas com a letra n, de normalizadas.

Tabela 2.5 - Símbolos das Predictoras Normalizadas

Símbolo	Variáveis Normalizadas
n1	Consumo
n2	Ano Situação
n3	Residencial
n4	Comercial
n5	Peso Apontamento
n6	Média 12 Faturamentos
n7	Score
y	Parecer

Fonte: Elaborado pelo autor (2014)

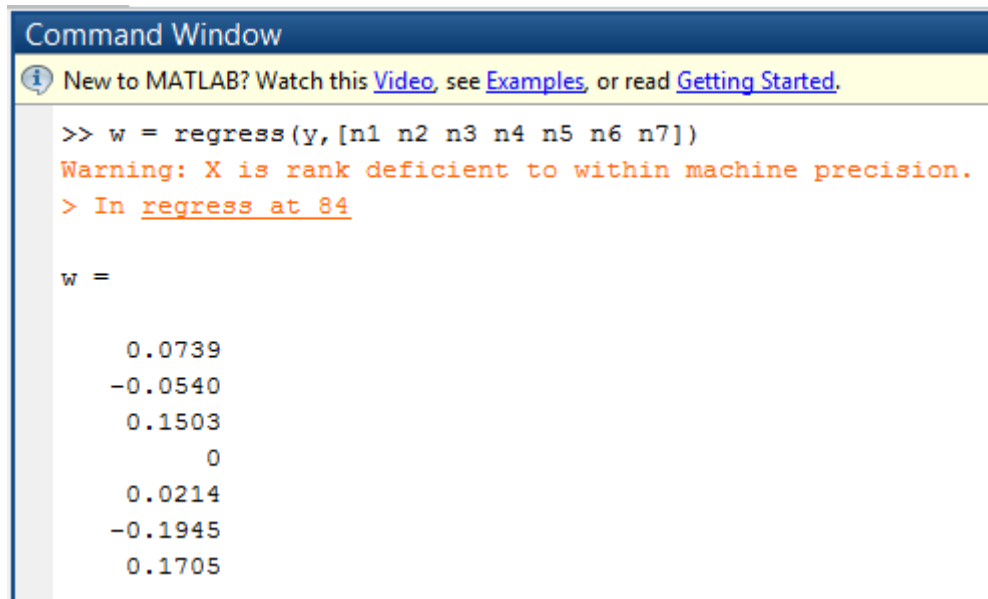
Aplicando o comando **regress** sobre as predictoras e a saída, são encontrados os pesos para cada variável, conforme tabela 2.5.

Tabela 2.6 - Pesos das Predictoras Determinados por meio do comando Regress

Símbolo	Variável	Pesos
n1	Consumo	0.0739
n2	Ano Situação	-0.054
n3	Residencial	0.1503
n4	Comercial	0
n5	Peso Apontamento	0.0214
n6	Média 12 Faturamentos	-0.1945
n7	Score	0.1705

Fonte: Elaborada pelo autor (2014)

Figura 2.24- Comando regress determinando peso das Preditoras



```

Command Window
New to MATLAB? Watch this Video, see Examples, or read Getting Started.

>> w = regress(y,[n1 n2 n3 n4 n5 n6 n7])
Warning: X is rank deficient to within machine precision.
> In regress at 84

w =

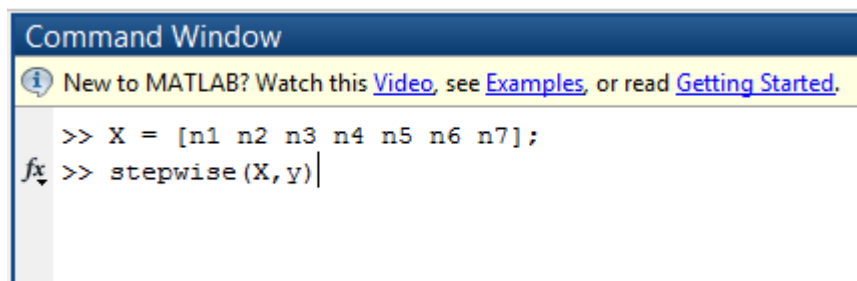
    0.0739
   -0.0540
    0.1503
         0
    0.0214
   -0.1945
    0.1705
  
```

Fonte: Matlab® (2014)

Os pesos demonstraram o nível de importância de cada variável, sendo que as preditoras n3 e n7 foram as mais relevantes na obtenção da resposta desejada, classe residencial e score respectivamente.

É possível determinar quais preditoras são realmente importantes, dentre as disponíveis. Para isso, o comando **stepwise** é utilizado. A figura 2.25 mostra o código de comandos no Matlab®, onde y é a saída do sistema.

Figura 2.25 - Comando Stepwise



```

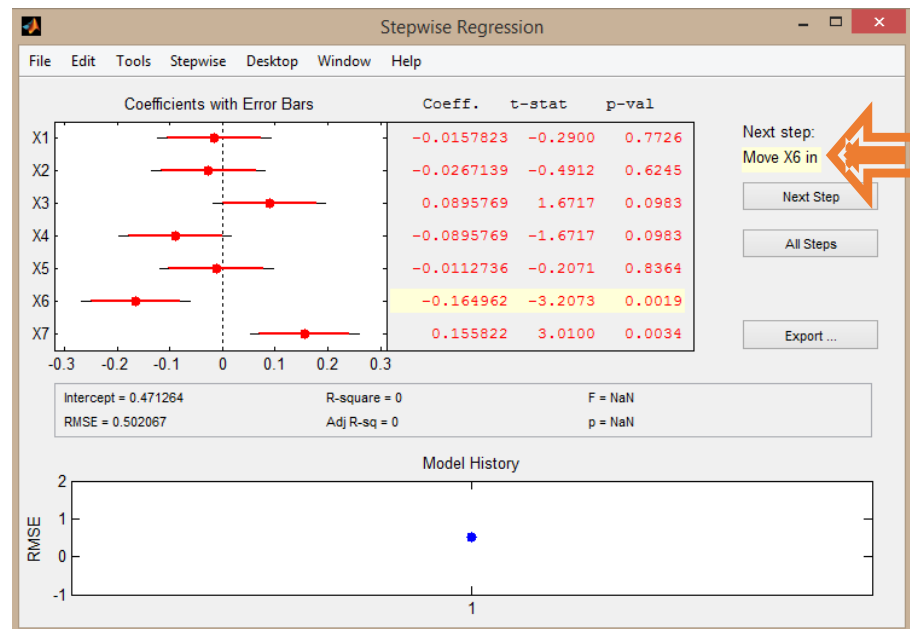
Command Window
New to MATLAB? Watch this Video, see Examples, or read Getting Started.

>> X = [n1 n2 n3 n4 n5 n6 n7];
fx >> stepwise(X,y)
  
```

Fonte: Matlab® (2014)

Após a execução do comando **stepwise** a janela da figura 2.26 aparece na tela do operador. O próximo passo sugerido pelo programa aparece ao lado direito da tela e indica que a preditora X6 deve ser movida.

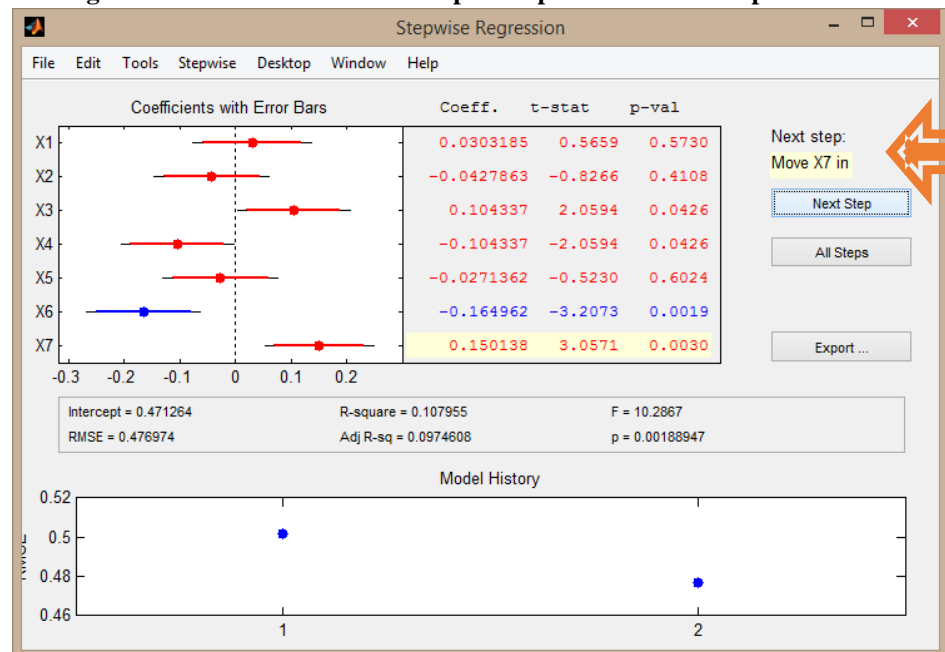
Figura 2.26 - Stepwise Regression



Fonte: Matlab® - Adaptado pelo Autor(2014)

Após mover a preditora X6, o próximo passo sugerido pelo programa é mover a preditora X7

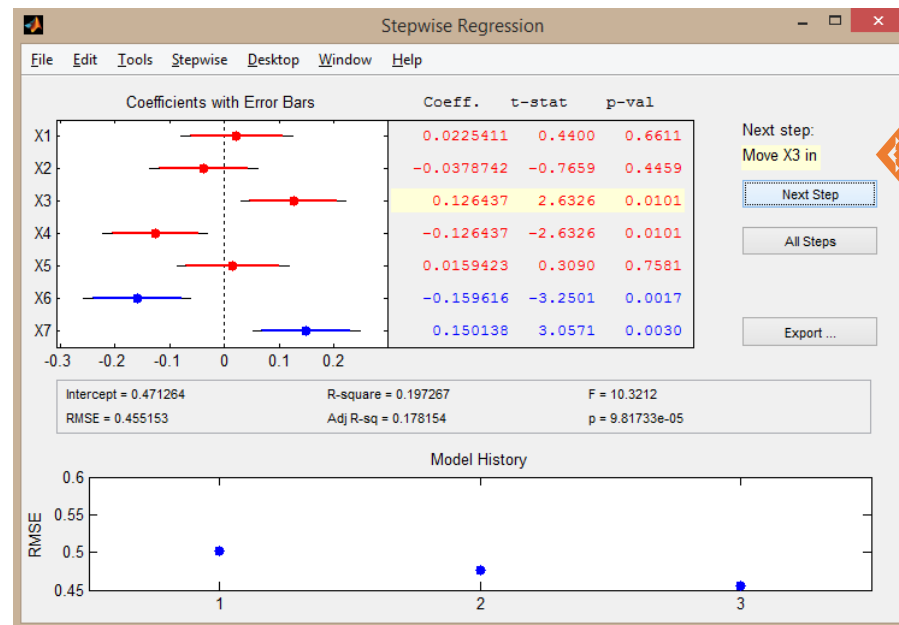
Figura 2.27 - Tela do comando Stepwise após movimento da preditora X6



Fonte: Matlab® (2014)

Após mover a preditora X7, o programa indica a necessidade de mover a preditora X3.

Figura 2.28 - Tela do comando Stepwise após movimento da preditora X7



Fonte: Matlab® (2014)

Após mover a preditora X3, o programa finaliza os passos e mostra as preditoras relevantes na obtenção da saída desejada. Elas aparecem em azul, no caso N3, N6 e N7

Figura 2.299 - Tela do comando Stepwise após movimento da preditora X3



Fonte: Matlab® (2014)

2.5.4 Conclusões

Após seguir todos os passos, o resultado obtido revelou que as preditoras X3, X6 e X7 equivalentes a N3, N6 e N7 são as variáveis que melhor representam a saída desejada.

Quando se utiliza muitas informações em bancos de dados, é interessante restringir o número de variáveis utilizadas, devido ao esforço computacional e dificuldades na manipulação dos dados. A ferramenta stepwise, determina as preditoras mais importantes para obtenção da resposta. Juntamente com o comando regress é possível obter os pesos de cada variável que geram o menor erro no valor da resposta de saída.

Portanto, as preditoras que devem ser utilizadas para determinação do perfil fraudador são as variáveis: Residencial, Média dos Doze Faturamentos e Score.

2.5.5 Análise dos Resultados

De uma amostra de 87 unidades consumidoras, três serão analisadas para a verificação da eficiência do método DNA da Fraude. Foram escolhidas unidades que tiveram três valores diferentes de score e que apresentaram ou não apontamento do leiturista. Será analisada a confiabilidade do resultado obtido por meio do score e a assertividade do seu valor para definição do perfil fraudador.

CASO 1) Unidade consumidora com suspeita de fraude, resultado do score igual a 225, com apontamento do leiturista e queda aparente no consumo. A tabela 2.6 define as características principais da UC 1.

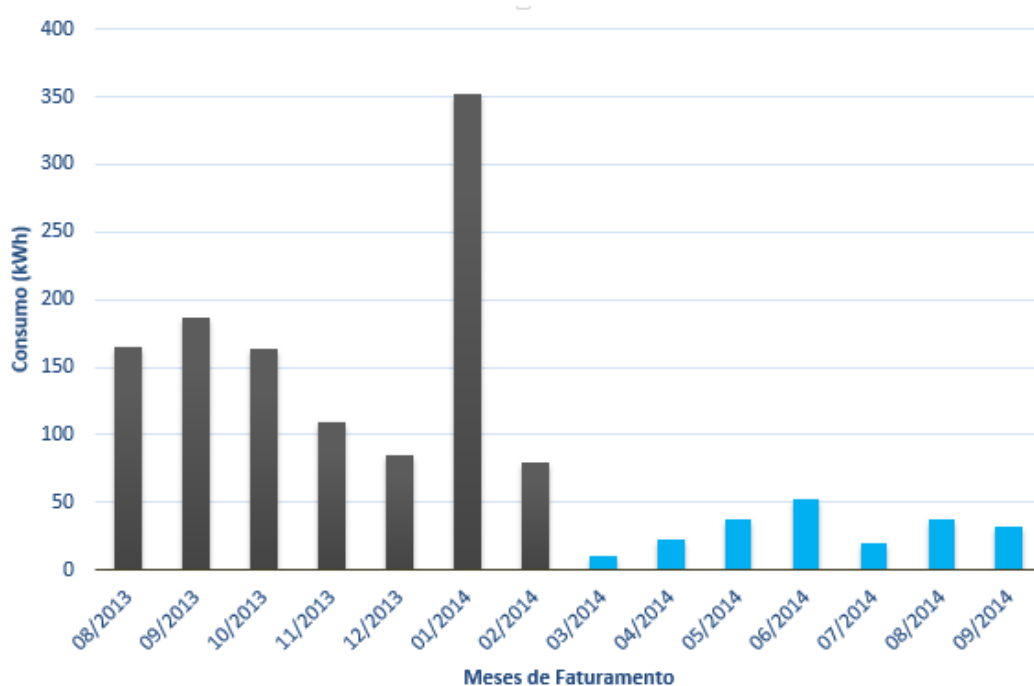
Tabela 2.7 - Características UC 1

Características	UC 1
Classe Principal	Residencial
Tipo de Fornecimento	Monofásico
Data da Ligação	2014
Fiscalização	Outubro/2014
Apontamento Leiturista	Suspeita de Fraude
Score	225

Fonte: Elaborado pelo autor (2014)

O histórico de consumo da UC 1 é retratado no gráfico 2.7. É possível notar a queda no consumo a partir do mês de março de 2014. Após ser realizada a fiscalização, a irregularidade foi constatada.

Gráfico 2.7 - Histórico de Consumo Anual - UC1



Fonte: Elaborado pelo autor (2014)

O parecer técnico apontou, que o neutro do ramal de entrada estava fora do borne do medidor, alterando o registro de consumo.

O score obtido e o apontamento do leiturista indicaram a irregularidade. A fiscalização em campo constatou que as duas variáveis estavam corretas em seus apontamentos. O trabalho conjunto das equipes de campo e da ferramenta computacional é fundamental para detecção das irregularidades e aumento da assertividade das fiscalizações.

CASO 2) Unidade consumidora com suspeita de fraude, com valor do score igual a 312, sem apontamento do leiturista e queda no consumo. A tabela 2.7 define as principais características da UC 2.

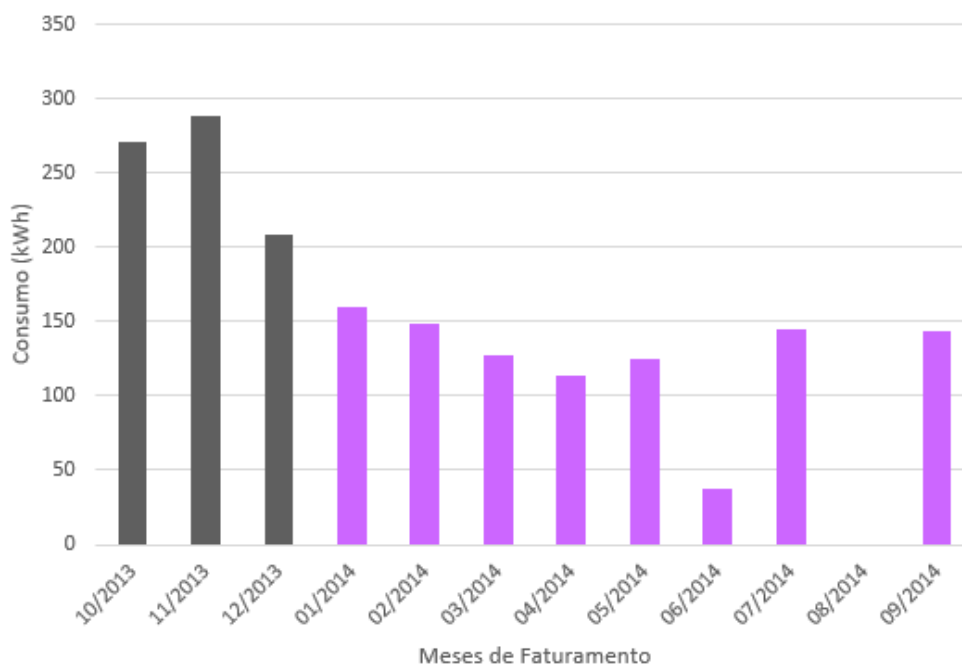
Tabela 2.8 - Características da UC2

Características	UC 1
Classe Principal	Residencial
Tipo de Fornecimento	Monofásico
Data da Ligação	2014
Fiscalização	Outubro/2014
Apontamento Leiturista	-
Score	312

Fonte: Elaborado pelo autor (2014)

O histórico de consumo da unidade consumidora 2 é retratado no gráfico 2.8. É possível notar a queda no consumo a partir do mês de janeiro de 2014. Após ser realizada a fiscalização, a irregularidade foi constatada.

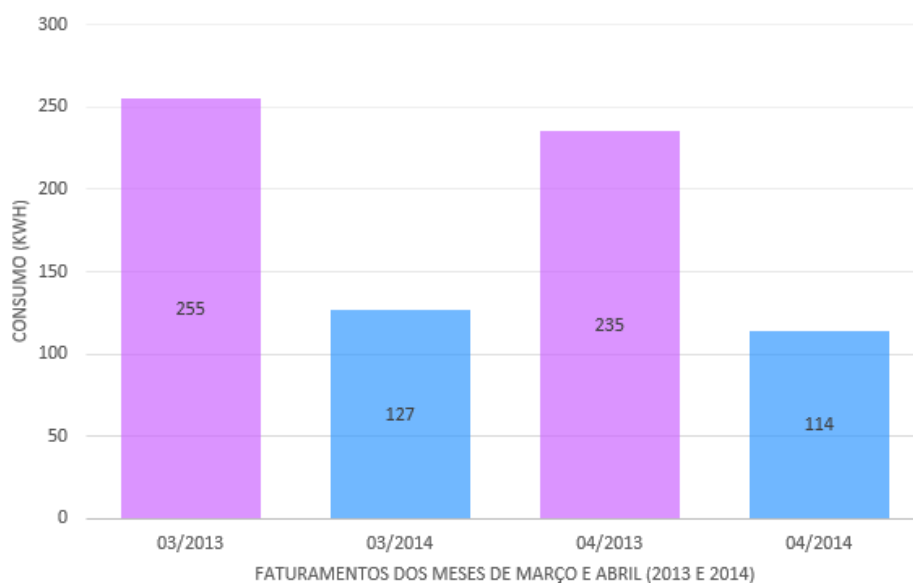
Gráfico 2.8 - Histórico de Consumo Anual – UC2



Fonte: Elaborado pelo autor (2014)

Uma outra análise que pode ser feita em conjunto com o resultado do score, é a comparação dos consumos da unidade consumidora do ano atual, com os mesmos meses em relação ao ano anterior. O gráfico 2.9 mostra a comparação dos consumos dos meses de março e abril de 2014 com os mesmos meses do ano de 2013.

Gráfico 2.9 - Histórico de Consumo referente aos meses de Março e Abril (2013 e 2014) – UC2



Fonte: Elaborado pelo autor (2014)

O gráfico 2.9 mostra um comparativo dos consumos dos meses de março e abril, nos anos de 2013 e 2014. A queda no perfil de consumo entre os dois anos é acentuada. Como não houve troca de titularidade, os hábitos de consumo tendem a permanecer iguais. Portanto, a diminuição no consumo pode configurar uma irregularidade na medição.

CASO 3) Unidade consumidora com suspeita de fraude, com valor do score igual a 245, com apontamento do leiturista suspeita de fraude, sem queda no consumo.

Mesmo com valor do score característico de fraude, este é um caso de unidade consumidora com parecer normal, ou seja um caso em que o apontamento feito pelo DNA e pelo leiturista falhou. A tabela 2.8 define as principais características da UC 3.

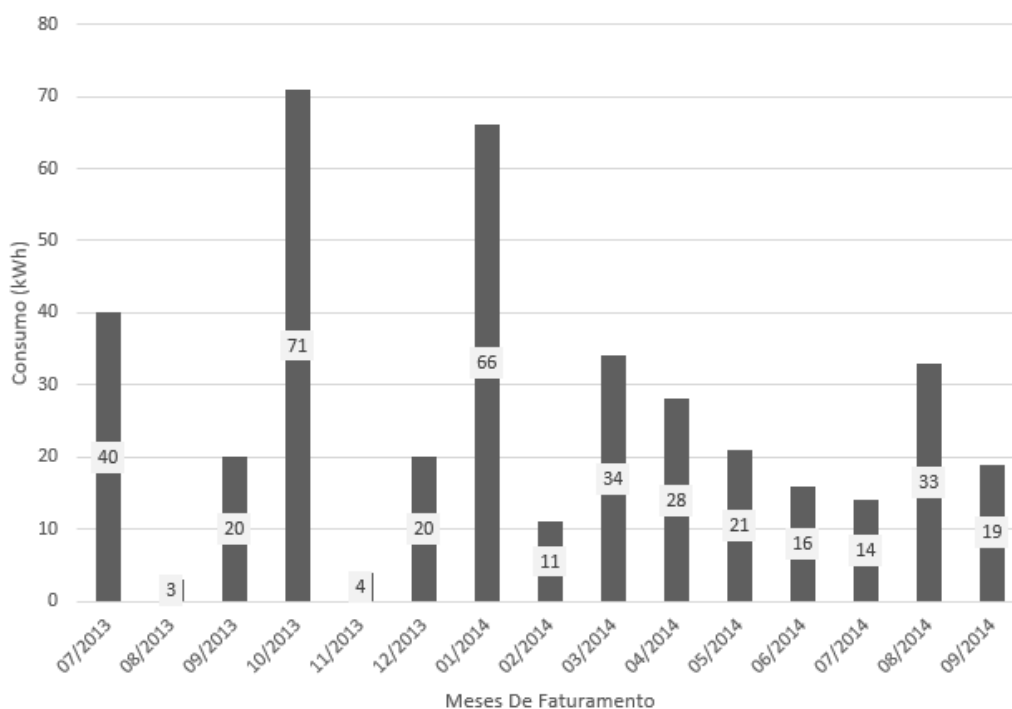
Tabela 2.9 - Características da UC3

Características	UC 1
Classe Principal	Residencial
Tipo de Fornecimento	Monofásico
Data da Ligação	2013
Fiscalização	Outubro/2014
Apontamento Leiturista	Suspeita de Fraude
Score	245

Fonte: Elaborado pelo autor (2014)

Não há queda de consumo visível no histórico da unidade consumidora 3. Existem alguns picos e quedas, mas não é possível identificar variação do consumo. Este cliente apresenta grande sazonalidade de consumo.

Gráfico 2.10 - Histórico de Consumo Anual – UC3



Fonte: Elaborado pelo autor (2014)

CAPÍTULO 3 - CONCLUSÕES

O algoritmo DNA da Fraude é uma ferramenta importante utilizada pela concessionária Enersul na detecção de fraudes. O principal objetivo desse sistema é aumentar a taxa de assertividade das fiscalizações e reduzir o número de inspeções. As ordens de serviço (fiscalizações) geradas são direcionadas em pontos estratégicos, em localidades com scores altos, em classes industriais e comerciais, nas classes residenciais bifásicas e trifásicas e em pontos onde as perdas do alimentador e transformador são altas.

A lógica empregada neste algoritmo é decodificar os dados e compactá-los em camadas, com o objetivo de simplificar o processo e torná-lo mais eficiente. Desta forma, o resultado do DNA é um somatório de valores decodificados dispostos em camadas. Cada variável que compõem as cinco camadas recebe pesos atribuídos heurísticamente, graças a experiência do operador e das equipes de campo

Como não há registros de como foi criado o algoritmo, a hipótese considerada foi que para cada camada é atribuído um peso específico e que estes pesos foram determinados por meio de uma regressão linear simples. Os dados disponibilizados pela empresa são bastante restritos e por isso não foi possível encontrar os valores reais dos pesos do algoritmo do DNA, então uma análise simplificada de regressão linear foi feita, apenas para demonstrar a utilização do processo. O software Matlab foi utilizado para determinação dos parâmetros.

O modelo apresentado não é adaptativo, portanto os pesos uma vez determinados nunca se alteram. Este é um ponto negativo, pois os tipos de fraudes de energia elétrica não permanecem estagnados, frequentemente aparecem novas irregularidades havendo a necessidade de reanálise das preditoras mais eficientes.

Uma das muitas soluções que pode ser empregadas na detecção de irregularidades é a implantação de redes neurais artificiais. Uma característica das RNAs é a capacidade de aprendizado e correção automática de erros, por meio do aprendizado supervisionado. Neste processo, existe um professor externo que estimula as entradas da rede por meio de padrões e supervisiona a saída calculada. O supervisor faz a comparação entre a saída encontrada e o valor da saída desejado com o objetivo de minimizar a diferença entre estes valores, os pesos são ajustados de forma a aproximar as respostas das duas saídas [11].

Uma solução mais simplista e não menos importante é o constante monitoramento dos pesos de todas as variáveis. Mesmo os pesos determinados heurísticamente devem ser monitorados e devem ter seus valores encontrados por meio de regressões lineares simples, já que o método estatístico é mais eficiente que o modelo empírico.

O último ponto de melhoria é a desfragmentação das cinco camadas, em várias outras. Muitas variáveis que não são correlacionadas estão dispostas na mesma camada, reduzindo a eficiência do algoritmo e aumentando a probabilidade de erros.

As perdas comerciais estão ligadas a vários fatores sociais e econômicos e o seu aumento retrata as condições precárias da população (inadimplência, violência, desemprego). No entanto, o segmento baixa renda é responsável por uma pequena parcela das perdas comerciais. Os grandes consumidores e a população de classe média ainda são os principais fraudadores, por isso é preciso considerar as características destes clientes e planejar medidas estratégicas para otimização e combate das perdas comerciais.

CAPÍTULO 4 - REFERÊNCIAS BIBLIOGRÁFICAS

[1] ANEEL, Agência Nacional de Energia Elétrica. Módulo 2: Revisão Tarifária Periódica das Concessionárias de Distribuição. Submódulo 2.6 – Perdas de Energia, v.1 p.4, 2013.

[2] GHOSH, S.; REILLY, D. L. Credit card fraud detection with a neural network. System Sciences, 1994. Informations Systems: Decision Support and Knowledge Bases System, Proceedings of the Twenty- Seventh Hawaii International Conference, v.3, p.621-630, 1994.

[3] TANIGUCHI, M.; HAFT, M.; HOLLMER, J.; TRESP, V. Fraud detection in communication networks using neural and probabilistic methods. Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference, v.2 p.1241-1244, 1998.

[4] REDNER, R. A.; WALKER, H. F. Mixture densities, maximum likelihood and the EM algorithm. SIAM Review, 1984.

[5] CABRAL, J. E.; PINTO, J. O. P.; GONTIJO, E. M.; REIS, J. Fraud detection in electrical energy consumers using rough sets. In: 2004 IEEE International Conference on Systems, Man, and Cybernetics; 2004.

[6] REIS, J.; GONTIJO, E. M.; MAZINA, E.; CABRAL, J. E.; PINTO, J. O. P. Fraud identification in electricity company customers using decision tree. In: 2004 IEEE International Conference on Systems, Man, and Cybernetics, p. 3730-3734, 2004.

[7] CELESC, Portal Grandes Clientes. Fator de Multiplicação, 2014. Disponível em portal.celesc.com.br/portal/grandesclientes/index.php?option=com_content&task=view&id=67&Itemid=97 . Acesso em dezembro de 2014

[8] ONS, Operador Nacional do Sistema Elétrico. Instalação do Sistema de Medição para Faturamento. Submódulo 12.2 - Procedimentos de Rede, 2004.

[9] ABRADEE. Associação Brasileira dos Distribuidores de Energia Elétrica. Disponível em www.abradee.org.br Acesso em novembro de 2014.

[10] BRAGA, A. P.; CARVALHO, A. P.; LUDERMIR, T. B. Redes Neurais Artificiais - Teorias e Aplicações. Rio de Janeiro : LTC, 2012. ISBN 978-85-216-1564.

[11] ZOU, K. H., TUNCALI K., STUART G. Statistical Concepts Series - Correlation and Simple Linear. From the Department of Radiology Brigham and Women's Hospital and Department of Health, Harvard Medical School. Boston, 2011