

Sistema de caracterização de perdas não técnicas baseado no perfil de consumidores para otimização de inspeções e recuperação de receita

Carlos Alberto Fróes Lima, Renato Aguiar Gurgel, João Paulo Ribeiro, Marley Apolinário Saraiva, Glauber Renato Colnago, Bernardo Marega Luz

Resumo - No processo de análise das perdas não técnicas de uma concessionária é importante a visão sistêmica do volume, histórico e da qualificação dos dados. Informações relevantes se apresentam como dependentes da realidade operacional, da estrutura de atendimento e consumo regionais. Esta observação não é óbvia e se torna mais evidente no momento de determinar indicadores de sucesso para a caracterização de perdas, redução de custos com inspeções e na efetividade das equipes de campo e de analistas de negócio. Fórmulas, algoritmos e resultados de pesquisas já realizadas podem ser somente referência de sucesso, sem garantir aplicabilidade direta devido às peculiaridades regionais, casos especiais, situações sociais e do negócio de energia distintos. Para a concessionária Coelce, com um grande número de clientes de baixo consumo e baixa renda, são geradas neste trabalho, abordagens para o direcionamento estratégico e efetividade na pesquisa por situações de perdas.

Palavras-chave – Inspeções, Perdas não técnicas, Perfil de consumo, Recuperação de receitas, Redes Neurais Artificiais

I. INTRODUÇÃO

Partiu-se neste projeto do entendimento dos procedimentos de caracterização de perdas realizados pela concessionária para os clientes do grupo B, com medidores convencionais, para criar uma sistemática diferencial que permitisse:

- Verificar a influência do histórico dos eventos e a utilização de regras na formulação das questões de negócio relacionadas a casos de estudos de elementos fraudadores e de elementos de evasão de receita na distribuição de energia elétrica;
- Criar o perfil característico de consumidores e de tendências de fraude e/ou perdas não técnicas;

Este trabalho foi desenvolvido no âmbito do Programa de Pesquisa e Desenvolvimento Tecnológico do Setor de Energia Elétrica regulado pela ANEEL e consta dos Anais do VI Congresso de Inovação Tecnológica em Energia Elétrica (VI CITENEL), realizado em Fortaleza/CE, no período de 17 a 19 de agosto de 2011.

Carlos Alberto Fróes Lima, Marley Apolinário Saraiva, Glauber Renato Colnago e Bernardo Marega Luz trabalham na KNBS Telecomunicações e Informática Ltda. (e-mail de contato: froes@knbs.com.br – www.knbs.com.br)

Renato Aguiar Gurgel e João Paulo Ribeiro trabalham na coelce (e-mail de contato: renato@coelce.com.br – www.coelce.com.br).

- Aplicar as regras de negócios existentes para a caracterização das inspeções realizadas e convergência para a sua formalização e adequação ao ambiente complexo das inúmeras variáveis verificadas;
- Definir regras e modelos novos, adequando e evoluindo a sistemática aplicada para a realização de inspeções na concessionária e efetividade das equipes em campo.

Foi fundamental inicialmente, o questionamento dos processos de compilação, reorganização, consolidação e consistência de dados, buscando a construção de um *datawarehouse* dedicado ao processo de análise de perdas não técnicas, organizando o conhecimento de forma a garantir a busca direta e independente dos diversos sistemas da concessionária.

Buscou-se, devido a complexidade e custo computacional da avaliação analítica proposta, garantir a independência do ambiente de produção da concessionária. Com este foco, foi estruturado um projeto de servidor de dados dedicado aos analistas de negócios para a busca de correlações, indicando melhorias para o desempenho operacional e permitindo visibilidade e aplicação de novas regras e correlações sem interferências sobre a produção atual da concessionária (considerando os limites naturalmente impostos pelas áreas de TI das empresas quanto a pesquisas de dados sistemáticas sobre seus grandes volumes de dados durante a operação diária).

Buscou-se efetivamente a definição de processos de apoio à decisão, permitindo:

- Investigação da correlação entre dados de faturamento existentes;
- Desenvolvimento das regras computacionais, formalização e auditoria do processo operacional implantado para a definição das inspeções;
- Avaliação temporal de correlação de consumo e relacionamento efetivado;
- Escolha de regiões e classes de consumo que deveriam ser priorizadas nos testes de hipótese desenvolvidos, bem como a verificação e adequação de regras existentes;
- Criação de indicadores de sucesso nas iniciativas de auditorias e inspeções;
- Identificação de focos de perdas e fraudes e propostas de ações para prevenção e correção.

Assim, foram estabelecidas condições e procedimentos metodológicos e gerado um sistema analítico de gestão das bases de dados de consumidores da concessionária, acoplado ao ferramental operacional existente e em produção de equipe de analistas. Buscou-se o conhecimento e ações para a redução de custos operacionais tanto na realização e sistemática de inspeções como na eficácia e assertividade do processo.

Com este objetivo, fez-se a extração de dados para análise do negócio e das situações de relacionamento com o cliente final, criando uma base de informação que deve ser mantida consistida e qualificada (como um processo constantemente executado pela equipe de analistas), incorporada na rotina operacional da concessionária. Possibilita, adicionalmente, de forma modular, ampliar-se a visão inicialmente estabelecida, como por exemplo, uma futura atuação com dados de bases censitárias e outras de correlação técnica, geoestatística e de ordens de serviço colocadas.

Foram realizados estudos e avaliação dos processos de faturamento e a proposição e qualificação de amostragens estatísticas para o reconhecimento das perdas não técnicas efetivas.

II. METODOLOGIA ADOTADA

O projeto foi desenvolvido considerando os seguintes procedimentos:

- Inicialmente foi realizado um estudo detalhado sobre o ambiente operacional da concessionária, foram reconhecidas as ferramentas em uso e as técnicas aplicadas com este ferramental. Foram também validados os processos existentes, o reconhecimento teórico existente sobre o assunto na equipe de trabalho. A partir deste ponto, foram realizados estudos de possibilidades para a realidade da concessionária, para seu ambiente de análise e produção. Os resultados deste estudo foram objeto de uma série de apresentações com a equipe da concessionária, que nortearam o desenvolvimento do projeto. No processo além da documentação disponível e de reuniões técnico-científicas, foram coordenadas e realizadas visitas da equipe de perdas à outra concessionária, a partir da indicação de necessidade de melhoria de ferramental. Estas visitas tiveram o intuito de apresentar soluções, ferramental e técnicas para a futura modernização das ações da área. É importante salientar que um dos objetivos do projeto foi melhorar os procedimentos existentes.
- Na segunda fase do projeto, foi desenvolvido um sistema computacional, base do projeto. Este sistema foi desenvolvido em JAVA e interfaces de banco de dados SQL SERVER, além de scripts na ferramenta em uso na concessionária, JMP. O sistema computacional acessa dados de faturamento para a qualificação de perfis para inspeção, permitindo a obtenção de resultados parciais que agregam valor ao Conhecimento que a Concessionária tem do seu negócio.

Assim, a modelagem final foi organizada de forma multi-dimensional, fortalecendo-se com o uso de ferramentas analíticas facilitadoras para a exploração pelos usuários do sistema proposto, porém dentro das limitações apresentadas pelo ferramental e conhecimento da equipe da concessionária. Foram apresentadas opções de ferramentas analítico-preditivas de última geração, resultado também de uma pesquisa realizada no âmbito do projeto.

O conceito e os resultados foram aplicados diretamente sobre o controle de perdas e de inspeções da concessionária, permitindo localizar causas de perdas e fornecendo indicadores para ações de prevenção e correção. A informação foi organizada, com a construção de protótipos e de uma base de dados, com a inserção de funções de avaliação temporal, de tendências de consumo e do conhecimento do consumidor, com regras de negócio bem determinadas para a avaliação de fraudes.

Este projeto utilizou os conceitos de datawarehouse que considera a organização das informações por assunto, integração, dimensão temporal, não volatilidade, níveis de granularidade, agregação e metadados, praticadas no ambiente operacional da concessionária, caracterizando-se assim uma quebra de paradigma nas abordagens lineares realizadas normalmente.

O sistema proposto compilou o histórico dos clientes, incluindo consumo e pagamentos efetuados, para criar o perfil e tendência de perdas, a partir do conhecimento existente de inspeções realizadas e TO emitidos.

O processo de apoio a decisão permite ainda:

- Investigação da correlação entre dados de faturamento existentes;
- Desenvolvimento das regras computacionais, formalização e auditoria do processo operacional implantado para a definição das inspeções;
- Avaliação temporal de correlação de consumo e relacionamento efetivados;
- Escolha das regiões e das classes de consumo priorizadas nos testes de hipótese desenvolvidos; verificação e adequação de regras existentes;
- Criação de indicadores de sucesso nas iniciativas de auditorias e inspeções;
- Ampliação do conhecimento do consumidor, possibilitando ações de melhoria de relacionamento com este cliente e gerando indicadores para atendimento;
- Identificação de focos de perdas e propostas de ações para prevenção e correção;
- Identificação de focos de fraudes e propostas de ações para prevenção e correção.

Foram realizadas sessões de questionamentos técnicos internas no âmbito da concessionária para validação junto às áreas, do sistema e técnica desenvolvidos.

Foram elaborados manuais de projeto e de manutenção do sistema desenvolvido, bem como treinados os usuários potenciais para o acompanhamento do sistema sob teste e análise.

Para que os objetivos deste projeto fossem alcançados, as várias etapas foram rigorosamente cumpridas e seus resultados ampliados passo a passo, de forma comprovada nos relatórios, apresentações e documentos elaborados durante a pesquisa e o projeto.

III. DESENVOLVIMENTO

A. O conhecimento do espaço operacional de análise de perdas da Concessionária

Iniciou-se a organização do conhecimento da base de clientes do Grupo B (Resolução ANEEL 456) [1], buscando-se a caracterização e priorização de avaliações de perdas e recuperação de receitas. Nenhuma alteração significativa, entretanto, precisa ser implementada para a nova resolução de 2010 (Resolução ANEEL 414) [2].

Neste contexto, foi feito o levantamento do consumo dos clientes, apresentado nas Figuras 1 e 2.

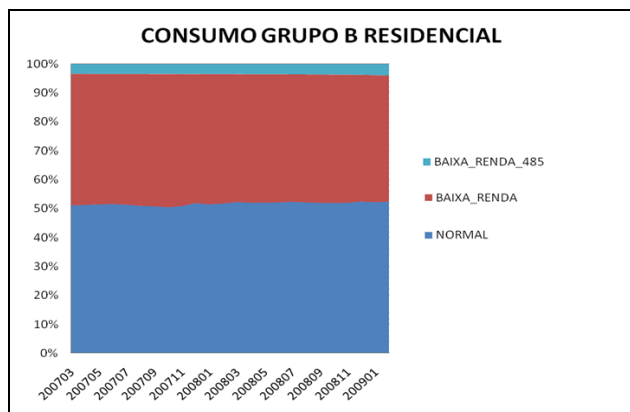


Figura 1- Percentual de consumo por classe

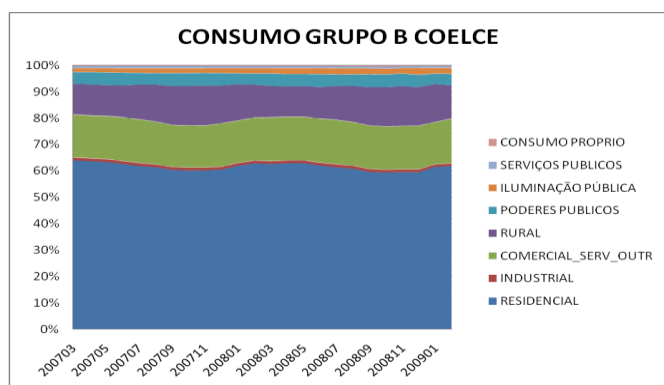


Figura 2 – Percentual de Consumo por subclasse na classe Residencial

Pode ser observada nestes gráficos toda a característica regional de distribuição de consumo, indicando a necessidade de particularização das análises a serem feitas. A aplicação de regras de caracterização de perdas teve também como pontos de atenção a organização operacional atual das equipes de perdas e inspeções.

Buscou-se neste estágio do projeto caracterizar as diversas ações de recuperação de receita, considerando a energia

registrada entregue e a quantidade de clientes, o conhecimento dos indicadores e a regionalização de ações de inspeções realizadas.

Neste momento de caracterização, foram feitas análises das bases de dados e consistência cadastral, necessárias para a validação da base de dados disponibilizada, resultando também em ações de recuperação de receita. Estas ações resultaram na detecção de algumas inconsistências no cadastro, expondo resultados operacionais e ações a serem tomadas.

Ressalta-se que o processo de organização do conhecimento com foco em perdas não-técnicas foi escolhido para a construção do modelo analítico proposto.

Algumas abordagens de mapeamento geográfico de consumidores foram buscadas, porém a não universalidade de dados de georeferenciamento da base da concessionária impediu a continuidade desta linha de pesquisas.

A concessionária, em sua estratégia logística de realização de inspeções divide suas ações estratégicas segundo suas equipes de campo, realizando historicamente uma divisão de ações apresentada na Figura 3.

Inspeções por região (FEV/2009)

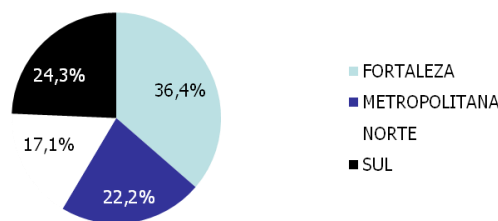


Figura 3 – Inspeções realizadas segundo logística da concessionária

Desta análise foi possível determinar em cada região as características de seus clientes. A partir dos resultados das inspeções realizadas, para cada região e para cada classe de clientes, foram determinadas a efetividade das ações tomadas e a possibilidade de ações futuras. Isto permitiu criar uma gama de indicadores, que nortearam o trabalho desenvolvido e a sistemática para as melhorias operacionais, tanto da equipe de analistas de negócio quanto para a estratégia das equipes de campo e de inspeções.

B. A consolidação de dados

A organização do processo de consolidação de dados e regras de negócio foi aplicada à base de dados disponibilizada pela empresa, priorizando-se as buscas segundo as regras de negócios trabalhadas. As dificuldades encontradas dizem respeito à dinâmica de aplicação de regras de negócio, de aplicações estatísticas e analíticas elaboradas não disponível no ferramental em uso pela equipe de analistas da concessionária. Com estas restrições, buscou-se fortalecer o entendimento tecnológico existente, evolutivo e uma reestrutur

turação com uma mudança cultural e operacional para um modelo orientado a banco de dados.

Antes da análise dos **volumes de energia recuperados após as inspeções** e da **efetividade das inspeções**, foi necessário definir essas duas grandezas, dentro da operação da concessionária.

Os consumidores que apresentam algum tipo de perda, geralmente, quando têm sua situação regularizada após a inspeção, possuem um aumento na medição de sua energia consumida média, salvo os casos em que o medidor media mais energia antes da inspeção. Assim, o **volume médio recuperado, considerado nas análises foi definido como a diferença entre a média dos consumos do segundo ao sexto mês após a inspeção e a média do consumo dos cinco meses anteriores à inspeção**. O mês seguinte à inspeção é descartado, pois pode haver grande variação do consumo antes de sua estabilização. A Figura 4 ilustra as médias dos consumos antes (MA) e depois (MD) da inspeção. O volume recuperado é, então, MD-MA.

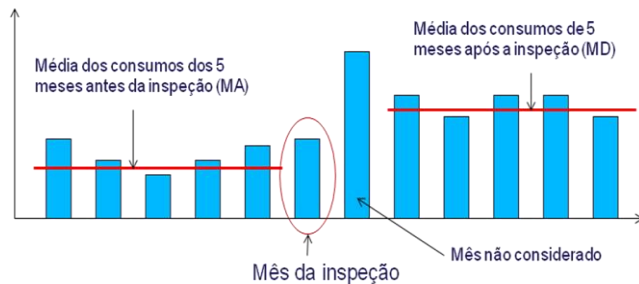


Figura 4 - Médias dos consumos antes e após as inspeções.

O valor esperado do volume recuperado é importante para distribuir o número de inspeções entre as classes. Este valor deve ponderar a efetividade das inspeções em uma determinada classe e o volume recuperado de energia dos consumidores daquela classe.

O objetivo inicial foi maximizar a efetividade nas inspeções. No entanto, alta taxa de efetividade não significa necessariamente maior recuperação de volume de energia. Exemplos deste caso são os consumidores da classe RESIDENCIAL BAIXA RENDA, que possuem consumos pequenos quando comparados às demais classes com uma alta incidência de situações que indicam perda, com questões sociais e culturais envolvidas.

O cálculo do volume de energia recuperado esperado é o produto da **taxa de efetividade nas inspeções (probabilidade de se encontrar um cliente com perda)**, pelo volume recuperado médio de cada consumidor. A Figura 5 mostra o volume recuperado médio ou esperado em cada classe. Com outras palavras, o gráfico apresenta o volume recuperado obtido com uma inspeção realizada a um consumidor ao acaso em uma determinada classe.

Esta análise proporcionou o questionamento inicial e o direcionamento para a caracterização de um procedimento efetivo de redes neurais para a sistematização das propostas de inspeções. Esta análise se baseou no histórico de efetividade na recuperação de receitas, na regionalização e logística existentes na concessionária. Não se propôs alterações,

neste momento, nesta logística como também não foram estruturadas outras ações de ampliação de subdivisão ou regionalização, respeitando-se as estratégias operacionais regionais aplicadas, bem como as dificuldades e organização vivenciados.

Ficaram evidenciadas nesta avaliação as características próprias dos consumidores de baixa renda na base da concessionária, que necessitam de um tratamento e uma política administrativa direcionada, estratégica, sendo que um tratamento somente de perdas não gerará resultados por si.

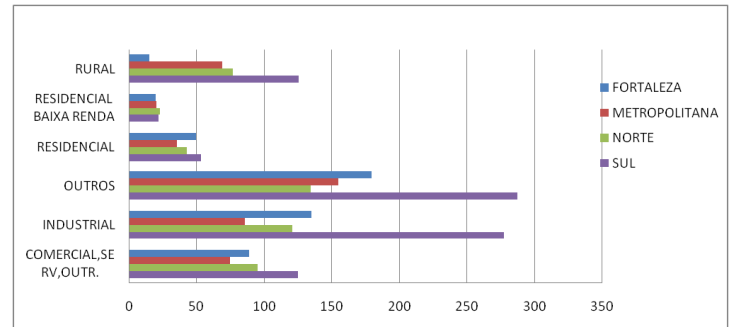


Figura 5 – Volume recuperado médio por consumidor de cada classe (KWh/mês), considerando 10 meses de análise

C. Proposição de um plano amostral

No processo de análise das perdas não técnicas da concessionária, torna-se relevante a visão sistêmica da qualificação da informação. Nesta visão está a proposição de avaliação por pesquisa e amostragem estatística¹ para validar números e a abordagem operacional da concessionária. Foi proposto e exercitado um procedimento estatístico baseado num plano amostral para obter uma visão diferenciada das perdas, avaliando o percentual de UC (unidades consumidoras) que geram perdas não técnicas do grupo B da COELCE, exceto os de baixa renda, por uma decisão estratégica de abordagem da base e do processo de perdas.

O Plano Amostral proposto inclui os seguintes aspectos: a determinação do tamanho da amostra, a seleção das UC, a forma de realização das inspeções das UC selecionadas e a geração de ordens de serviços especiais para sua execução, a validação das informações coletadas resultantes da atividade de inspeção (auditada) e o cálculo do percentual de UC que geram perdas não técnicas.

As Unidades Consumidoras devem ser selecionadas de forma aleatória para a avaliação amostral e assim foi feito. Além disto, recomendou-se a segregação em divisões, classes e número de fases segundo a instalação existente e a atuação estabelecida pela própria concessionária, além da observância da relação custo-benefício do processo de amostragem como um todo.

Com relação ao tamanho da amostra, sua definição foi estabelecida com base em métodos estatísticos considerando

¹ Maiores detalhes sobre técnicas de amostragem estatística podem ser vistos em **Erro! Fonte de referência não encontrada.**

uma margem de erro aceitável (em torno de 1%). A definição da margem de erro foi adotada em conjunto e em reunião com a equipe de análise de perdas da COELCE, ponderando-se as limitações operacionais (e de orçamento). Foi inicialmente sugerida, a partir da análise prévia e histórica das perdas detectadas nas inspeções realizadas, a inspeção de cerca de 5.320 (cinco mil trezentos e vinte) UC, para se garantir uma margem de erro de 1%.

Sugeriu-se adicionalmente, que as inspeções fossem auditadas. Considerando os procedimentos usuais da COELCE com o objetivo de garantir a precisão dos resultados, buscou-se focar a auditoria realizada nas UC da lista de amostragem, ampliando-se a segurança da qualidade dos resultados obtidos, objetivando a crítica dos seguintes dados:

- Temporais: dia e hora da execução;
- Cadastrais: empresa responsável pela execução, técnico que executou;
- Conteúdo de leituras;
- Relatórios: laudos da inspeção.

D. Modelagem de Redes Neurais Artificiais e evolução de resultados de inspeção

Foram utilizadas na modelagem de qualificação de unidades consumidoras para inspeção ferramentas de Redes Neurais Artificiais (RNA), buscando detectar consumidores da concessionária COELCE que apresentam perdas não técnicas. Partiu-se do interesse atual da concessionária de identificar um consumidor com perda como aquele que possa gerar como resultado da inspeção um termo de ocorrência (TO). Isto ocorre quando, após inspecionar um cliente, verifica-se que existe alguma condição que implique em fraude ou outro motivo de perda não técnica como, por exemplo, problema no medidor.

Na aplicação das RNA, os dados de consumidores foram divididos em dois grupos para as análises: um deles definido como conjunto de treinamento e outro como de validação. Estes conjuntos se referem a consumidores já inspecionados.

Considerando o comportamento diferente dos consumidores foram estabelecidos inicialmente 64 grupos de consumidores, obtidos através de segregações de classes e regiões logísticas da concessionária. Verificando-se as características comuns de comportamento apresentados destes 64 grupos e padrões comuns, estes foram reduzidos para 17 grupos após diversos testes de possíveis agregações realizadas buscando-se a generalização. É importante ressaltar que a divisão e agregação foram realizadas com o auxílio das Redes Neurais de Kohonen, um tipo de Rede Neural com aprendizado não supervisionado que divide um conjunto de dados em grupos [4].

E. Variáveis de entrada da RNA

Diversos tipos de dados foram testados na modelagem RNA na tentativa de encontrar e representar mais adequadamente o fato de um consumidor apresentar perda. Os da-

dos mais representativos geraram variáveis de análise, cuja relevância foi testada exaustivamente nas RNA estabelecidas:

a. Dados referentes ao padrão de consumo

Este tipo de dados visou detectar consumidores com perda de energia elétrica através de seu histórico de consumo. Os consumidores alvo destes dados são aqueles que apresentam variações de energia fora do comum do seu grupo.

b. Coeficiente de Variação

Este dado mede a variação de consumo de cada consumidor ao longo de um período de tempo. Ele é calculado como sendo a divisão entre o desvio padrão e a média dos 12 últimos consumos de uma UC, a partir do mês de referência.

c. Média Normalizada

A média normalizada tem como objetivo situar o percentual de consumo de um cliente com relação aos seus consumos máximo e mínimo. Este tipo de dado, juntamente com o coeficiente de variação, pode indicar picos de consumo, queda abrupta, ou comportamento mesmo regular de consumo ao longo do tempo.

Para calcular este dado, obtêm-se os valores máximo (C_{max}) e mínimo (C_{min}) dos 12 últimos consumos, que são, por sua vez, normalizados segundo a fórmula

$$C_i norm = \frac{C_i - C_{min}}{C_{max} - C_{min}}$$

Onde C_i representa o consumo de um cliente no mês i , sendo $i = 12$ o mês mais atual.

Por último, calcula-se a média dos valores $C_i norm$, para $i = 1, \dots, 12$, obtendo-se, desta forma, o consumo médio normalizado.

d. Consumo relativo à média de consumo da localidade

Com este dado pretendeu-se indicar o quanto o consumo médio de cada cliente varia em relação ao consumo médio de sua localidade. Cada localidade foi determinada pelos campos Lote e Local fornecidos pela COELCE. Cada consumidor possui um número de lote e um de local (agência), segundo a base de dados da concessionária. Supondo que existem n lotes e m locais, então existem m vezes n ($m*n$) localidades diferentes.

Matematicamente, considerando que \bar{C}_{LOC} é o consumo médio dos consumidores de uma localidade e \bar{C} é o consumo médio de um cliente da mesma localidade, tem-se que o consumo relativo para aquele cliente é

$$C_{RELAT} = \frac{\bar{C}_{LOC} - \bar{C}}{\bar{C}_{LOC}}$$

e. Dados referentes a históricos de irregularidade

Este tipo de análise busca detectar aqueles consumidores que já apresentaram ou apresentam algum indício de irregularidade.

f. Indicador de perda na localidade ou vizinhança

Esta variável indica o percentual de consumidores com perda em uma determinada localidade, ou seja, a quantidade de consumidores que apresentaram TO em uma localidade, dividida pela quantidade de consumidores da mesma localidade.

No cálculo deste dado também utilizou-se a localidade. Conta-se, para cada localidade, quantos consumidores tiveram algum termo de ocorrência (TO). Em seguida conta-se a quantidade de consumidores da localidade. A divisão entre quantos consumidores tiveram TO e o número de consumidores da localidade é o indicador de perda na vizinhança, ou localidade. Este indicador é percentual. No caso de possuir o valor 1 para uma determinada localidade, significa que 100% dos seus consumidores inspecionados deram origem a um termo de ocorrência. Para aquelas localidades que não tiveram nenhuma inspeção, aloca-se o valor 0 para seu indicador de perda na localidade.

Este tipo de dado foi utilizado porque, segundo [1], quando um consumidor comete fraude, existe probabilidade alta de consumidores da vizinhança também cometerem. Esta hipótese foi utilizada neste espaço de análise, considerando a sua aplicabilidade prática.

g. Dívida Atual

Este dado refere-se a dívida atual do consumidor com relação à Concessionária, em reais. Buscaram-se com esta variável, consumidores que não conseguiram pagar a conta de energia ou mesmo consumidores que estão descontentes com o preço da energia ou serviço, e por isto possam cometer fraude ou furtar energia.

h. Histórico de TO

Foi agregado um campo TO à base de análise, positivado para o consumidor que foi inspecionado em um determinado mês e que resultou em um termo de ocorrência. Segundo o estudo feito em [6], muitos consumidores fraudulentos são reincidentes e esta reincidência acontece em períodos curtos.

F. Resultados

Os exercícios das RNA foram feitos para cada um dos 17 grupos. Foram comparados os acertos verificados na prática de inspeções da concessionária e os acertos obtidos com a metodologia proposta. Para o melhor entendimento dos acertos, define-se o acerto obtido com as RNA:

$$(\text{Acerto 1})/(\text{Acerto 1} + \text{Erro 10})$$

onde **Acerto 1** é quantidade de vezes que a Rede Neural referente a um grupo classificou corretamente, no conjunto de validação, consumidores com perda. **Erro 10** é

a quantidade de vezes que a Rede Neural indicou um consumidor como possuindo perda quando o mesmo não possuía. A soma **Acerto 1 + Erro 10** indica a quantidade de consumidores que a Rede Neural indicou como apresentando perda.

O acerto Coelce é o percentual de consumidores que apresentaram perda (independente do resultado das RNA) com relação ao total de consumidores do conjunto de validação. A Tabela 1 apresenta as diferenças percentuais entre os dois tipos de acertos. Assim, supondo que o acerto da Coelce para algum grupo foi $x\%$, e o das RNA $y\%$, a tabela mostra a relação $(y\% - x\%)/x\%$.

Percebe-se na tabela que as diferenças os valores são sempre positivos, mostrando resultados melhores que os praticados na Coelce atualmente. A diferença mínima foi de 28,82% e a máxima de 519,07%. Agregando todos os grupos, tem-se que a diferença de acertos é de 157,45%.

Por outro lado tem-se o acerto de classificação da metodologia proposta. Este acerto pode ser escrito como:

$$(\text{Acerto 0} + \text{Acerto 1})/(\text{Acerto 0} + \text{Acerto 1} + \text{Erro 01} + \text{Erro 10}),$$

onde **Acerto 0** é a quantidade de vezes que a Rede Neural acertou ao classificar consumidores sem perda e **Erro 01** é a quantidade de erros quando a Rede Neural indicou um consumidor como não possuindo perda quando o correto era o inverso.

O acerto de classificação considerando todos os grupos foi de 87,18%.

Tabela 1. Comparação entre metodologia proposta e a prática atual da concessionária

GRUPO	Diferença percentual entre acertos RNA e COELCE	% Acerto classificação
1	166,71	89,14
2	244,68	92,91
3	519,07	92,33
4	152,74	88,48
5	220,50	90,88
6	28,82	79,51
7	213,94	85,33
8	111,42	79,71
9	80,00	84,48
10	92,08	82,71
11	134,89	81,64
12	113,73	69,49
13	308,73	86,64
14	99,73	67,29
15	257,51	88,36
16	58,43	69,80
17	127,30	85,56
Total	157,45	87,18

G. Sistema computacional

Na segunda fase do projeto, foi desenvolvido um sistema computacional, base do projeto. Este sistema foi desenvolvido em JAVA [7] e interfaces de banco de dados POSTGRESQL [8], além de *scripts* na ferramenta em uso na concessionária. O sistema computacional acessa dados de faturamento para a qualificação de perfis para inspeção, permitindo a obtenção de resultados parciais que agregam valor ao conhecimento que a Concessionária tem do seu negócio.

Assim, a modelagem final foi organizada de forma multi-dimensional, fortalecendo-se com o uso de ferramentas analíticas facilitadoras para a exploração pelos usuários do sistema proposto, no contexto das limitações apresentadas pelo ferramental e conhecimento da equipe da concessionária. Foram apresentadas opções de ferramentas analítico-preditivas de última geração, resultado também de uma pesquisa realizada no âmbito do projeto.

Foram utilizados os conceitos de *datawarehouse* que considera a organização das informações por assunto, integração, dimensão temporal, não volatilidade, níveis de granularidade, agregação e metadados, praticadas no ambiente operacional da concessionária, caracterizando-se assim uma quebra de paradigma nas abordagens lineares realizadas na produção atual.

O sistema proposto compilou o histórico dos clientes, incluindo consumo e pagamentos efetuados, para criar o perfil e tendência de perdas, a partir do conhecimento existente de inspeções realizadas e TO emitidos.

Vale ressaltar o curto espaço de tempo para a organização do conhecimento, da pesquisa operacional e da realização de testes efetivos de validação da metodologia, conceitos e sistemas. Um projeto de caracterização de perdas não técnicas deve ser desenvolvido em pelo menos dois anos, com o primeiro ano dedicado a organização operacional e o segundo ano no acultramento de técnicas, aplicação e refinamento de regras de negócio, de verificação operacional, sistematização e abrangência da solução. A equipe foi treinada, porém o acompanhamento e resultados mais abrangentes não puderam fazer parte do escopo no tempo estabelecido pela concessionária para esta pesquisa e execução. Este aprendizado é bastante relevante para as demais concessionárias e para uma continuidade procedural de ações.

IV. CONCLUSÕES

O projeto foi desenvolvido considerando o estudo detalhado sobre o ambiente operacional da concessionária, foram reconhecidas as ferramentas em uso e as técnicas aplicadas, com este ferramental. Foram também validados os processos existentes e o reconhecimento teórico da equipe de trabalho. A partir deste ponto, foram realizados estudos de possibilidades para a realidade da concessionária, para seu ambiente de análise e produção. Os resultados deste estudo foram objeto de uma série de apresentações com a equipe da concessionária, que nortearam o desenvolvimento final do projeto. No processo além da documentação dispo-

nível e de reuniões técnico-científicas, foram coordenadas e realizadas visitas da equipe de perdas à outra concessionária, a partir da indicação de necessidade de melhoria de ferramental. Estas visitas tiveram o intuito de apresentar soluções, ferramental e técnicas para a futura modernização das ações da área. É importante salientar, que um dos objetivos do projeto foi melhorar os procedimentos existentes.

O conceito e os resultados foram aplicados diretamente sobre o controle de perdas e de inspeções da concessionária, permitindo localizar causas de perdas e fornecendo indicadores para ações de prevenção e correção. A informação foi organizada, com a construção de protótipos e de uma base de dados, com a inserção de funções de avaliação temporal, de tendências de consumo e do conhecimento do consumidor, com regras de negócio bem determinadas para a avaliação de fraudes.

Nesta análise, buscou-se avaliar as características inerentes do negócio para o atendimento do grupo B, foco do projeto. A segmentação em classes, enfatizando-se a segregação dos clientes considerados como baixa renda, mostrou-se de extrema importância para o ambiente da concessionária. A Coelce, em particular, possui um percentual muito grande de clientes de baixa renda, com características próprias de consumo e retorno efetivo financeiro complexo, quando acionado para pagamento ou regularização. Os clientes residenciais representam cerca de 75,6% dos clientes em carteira, sendo que 54,1% são clientes de baixa renda!

Assim, pondera-se e reafirma-se a necessidade de se separar formalmente o espaço de análise de clientes de baixa renda dos demais consumidores e a organização de uma abordagem político-administrativa especial para este grupo de clientes, com estratégias sociais inerentes ao resgate destes clientes para a base operacional, adimplente e geradora de receitas. Não foi foco deste trabalho uma análise amostral para os clientes de baixa renda, embora toda a técnica e metodologia aplicada possam ser estendidas para este segmento.

Buscou-se, através de ferramentas de Redes Neurais Artificiais (RNA), detectar consumidores da concessionária COELCE que apresentam perdas não técnicas, dentro do grupo B. Para isto, foi necessário, primeiramente, organizar e definir com a concessionária um consumidor com perda, associando-o a um termo de ocorrência - indicação explícita de atuação ou ação, verificando-se alguma condição que implique em perdas não técnicas, seja ela fraude ou outro motivo não intencional como, por exemplo, problema no medidor.

Nesta análise, evidenciou-se a segmentação natural da base e do negócio de forma geográfica, seguindo as dificuldades e a realidade apresentada pelas equipes de inspeção, estruturadas para atendimento segundo uma demanda explícita, regionalizada e dependente de custos diferenciados.

Foi então direcionada a análise pela base de dados tabular, construindo-se um *datawarehouse* de referência para a pesquisa. Foi construído um sistema no ambiente de labora-

tório, utilizando ferramentas não técnicas e acadêmicas para esta modelagem. Este ambiente não pode ser reproduzido no espaço atual de produção da concessionária, pelas limitações de ferramental. Assim, foi desenvolvido um subconjunto que expressasse a metodologia e os resultados e criado um ambiente de produção organizado, novo e implementável para a produção na concessionária. Este ambiente foi transferido e operacionalizado.

Os resultados obtidos representam um novo questionamento para a concessionária em sua análise de perdas, segmentação e forma de atuação.

V. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] ANEEL - Resolução Normativa Nº 456, de 29/11/2000
- [2] ANEEL - Resolução Normativa Nº 414, de 9/9/2010
- [3] COCHRAN, William Gemmell. Sampling techniques. 3rd. ed. New York: J. Wiley, c1977. xvi, 428p.
- [4] Monedero, Íñigo; Biscarri, Félix; Leon, Carlos; Biscarri, Jesús; Millán, Rocio. *MIDAS: Detection of Non-technical Losses in Electrical Consumption Using Neural Networks and Statistical Techniques. Computational Science and Its Applications - ICCSA 2006.*
- [5] Penin, Carlos A. S. Combate, *Prevenção e Otimização das Perdas Comerciais de Energia Elétrica.* São Paulo, 2008. Tese de Doutorado – Universidade de São Paulo.
- [6] Regueira, Bruno; Souza, Nivaldo P. F.; Ferreira, Fabiano C. *Experiências e Ações no Combate a Perdas Comerciais com Resultados na Otimização de Perdas Técnicas.* Anais do XVII Seminário Nacional de Distribuição de Energia Elétrica. Belo Horizonte, 21 a 25 de Agosto de 2006.
- [7] Java™ Runtime Environment 6 Update 1, disponível em <http://www.java.com> , acessado em 18/01/2010.
- [8] PostgreSQL 8.2, disponível em <http://www.postgresql.org> , acessado em 18/01/2010.