

Detecção de Perdas Não Técnicas via Rede Neural ARTMAP *Fuzzy* em Sistemas de Distribuição de Energia Elétrica

Vitor G. Silveira*, Andréia S. Santos**, Mara L.M. Lopes**, José F. Resende da Silva*, Lucas Teles Faria*

*UNESP – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Rosana
Av. dos Barrageiros, 1881, 19274-000, Rosana, São Paulo, Brasil

** UNESP – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de Ilha Solteira
Av. Brasil, 56, 15385-000, Ilha Solteira, São Paulo, Brasil

E-mails: vitor.silveira@unesp.br, andreia.faria@unesp.br, mara.lopes@unesp.br,
jose.resende@unesp.br, lucas.teles@unesp.br

Abstract: Non-technical losses (NTLs) or commercial losses are caused by multiple factors such as: energy theft, energy meter fraud, self-reconnection, damaged energy meter, consumer default and others. NTLs cause significant financial losses to power utilities, undue changes in the distribution network and damage to consumer units (CUs) with the increase in their energy bill. In this context, we propose a methodology based on artificial neural networks (ANNs) to detect NTLs in power distribution network. The methodology has two modules: (i) extraction of statistical attributes and (ii) module for classification based on the Fuzzy ARTMAP. The proposed methodology presents high indexes for the metrics: success rate (99.0%), reliability (93.6%), and specificity (96.7%). Therefore, it has good coverage for detecting irregular CUs and a high success rate in field inspections. Thus, costly field visits by inspection teams in regular UCs are drastically reduced.

Resumo: As perdas não técnicas (PNTs) ou perdas comerciais são ocasionadas por múltiplos fatores como: furtos de energia via ligações irregulares ou clandestinas, fraudes embutidas no medidor de energia, autorreligações, medidores avariados, inadimplência dos consumidores e outros. Essas perdas acarretam prejuízos financeiros significativos às distribuidoras de energia, à rede de distribuição e às unidades consumidoras (UCs). Nesse contexto, propõe-se neste estudo uma metodologia baseada em redes neurais artificiais (RNAs) para detecção de PNTs na rede de distribuição de energia elétrica. A metodologia possui dois módulos: (i) extração de atributos estatísticos baseados em regimes e (ii) módulo classificador baseado na RNA ARTMAP *Fuzzy*. O sistema classificador proposto apresenta altos índices para as métricas: taxa de sucesso ($TS = 99,0\%$), confiabilidade negativa ($CN = 93,6\%$) e especificidade ($ESP = 96,7\%$). Portanto, a metodologia proposta apresenta boa cobertura na detecção majoritária de UCs irregulares e alta taxa de sucesso nas inspeções em campo. Assim, reduzem-se as visitas indesejadas e dispendiosas das equipes de inspeção em UCs regulares.

Keywords: Artificial Neural Networks (ANNs); Commercial Losses; Electrical Distribution Systems; Energy Theft; Fuzzy ARTMAP; Non-technical Losses (NTLs).

Palavras-chaves: ARTMAP *Fuzzy*; Perdas Não Técnicas (PNTs); Perdas Comerciais; Furto de Energia; Redes Neurais Artificiais (RNAs); Sistema de Distribuição de Energia Elétrica (SDEE).

1. INTRODUÇÃO

As perdas não técnicas (PNTs) ou perdas comerciais acarretam prejuízos financeiros vultosos às distribuidoras de energia. Além disso, elas contribuem significativamente para aumento na fatura de energia de todas as unidades consumidoras (UCs), uma vez que uma parcela dos prejuízos devidos aos furtos e fraudes de energia são repassados às UCs (ANEEL, 2020). Adicionalmente, a manipulação da rede elétrica por pessoas não autorizadas traz um risco iminente de acidentes fatais e prejuízos à qualidade e a confiabilidade da rede elétrica. Portanto, o combate às PNTs proporciona maior segurança, recuperação de parte do faturamento das

distribuidoras e; conseqüentemente, redução na conta de energia elétrica.

Conforme Fig. 1, as perdas totais representaram cerca de 13,8% da energia injetada no sistema elétrico em 2019. Essa energia perdida é equivalente ao consumo de energia elétrica das regiões norte e centro-oeste no ano de 2018 (ANEEL, 2020). As PNTs estão presentes em todos os sistemas de distribuição e têm um nível elevado em muitas distribuidoras brasileiras, principalmente, naquelas pertencentes às regiões norte e nordeste (ABRADEE, 2018).

As perdas totais no sistema de distribuição de energia elétrica (SDEE) são definidas como sendo toda a energia

elétrica injetada na rede elétrica e que não é faturada juntos às UCs (ABRADEE, 2018).



Fig. 1. Perdas totais sobre a energia injetada em 2019 (ANEEL, 2020).

As perdas de energia elétrica no SDEE são classificadas em perdas técnicas e em PNTs. As perdas técnicas estão associadas às “perdas da engenharia”, como por exemplo, as perdas Joule nos equipamentos da rede elétrica (cabos, transformadores, religadores, etc.) durante o transporte de energia, desde o local de geração até o consumidor final (ANEEL, 2020).

Por outro lado, as PNTs são ocasionadas principalmente por furtos e fraudes de energia. O furto de energia consiste no desvio direto no medidor de energia. A ligação clandestina é uma modalidade de furto de energia, onde há uma conexão direta, a partir da rede de distribuição, para um consumidor ilegal e não inscrito na base de clientes da distribuidora. As fraudes ocorrem quando determinadas UCs realizam adulterações no sistema de medição da sua residência, comércio ou indústria de modo que paguem efetivamente por uma parcela menor do total de energia consumida, devido à fraude implantada (ABRADEE, 2018).

As distribuidoras de energia realizam inspeções periódicas em UCs com suspeita de irregularidades. Essas inspeções apresentam custo financeiro elevado, pois cada UC suspeita é avaliada minuciosamente; portanto, essas inspeções demandam muito tempo das equipes em campo – elevado custo homem-hora.

Nesse contexto, é necessário o uso de ferramentas computacionais eficientes para uma seleção aprimorada das UCs que devem ser inspecionadas. Em outras palavras, idealmente, todas as UCs inspecionadas deveriam possuir alguma irregularidade; caso contrário, as equipes de inspeções estariam “perdendo tempo” ao inspecionarem UCs regulares.

As inspeções são motivadas por denúncias dos leituristas e da população. Algumas distribuidoras utilizam o conhecimento empírico de seus técnicos que efetuam análises semiautomáticas em planilhas eletrônicas em busca de UCs com histórico de consumo mensal em kWh atípico e; portanto, candidatas a receberem inspeções. Outras distribuidoras minoritárias possuem sistemas computacionais sofisticados e baseados em técnicas de computação flexível ou *soft computing* que são capazes de extrair conhecimento oculto em uma “massa” de dados.

A seleção de UCs possivelmente irregulares e; portanto, alvos de inspeções, baseiam-se majoritariamente em técnicas de *soft computing*. Alguns exemplos da literatura especializada: redes neurais artificiais (RNAs) *perceptron* multicamadas (PMC) (Faria, Padilha-Feltrin e Minussi, 2012),

lógica *fuzzy* (Cabral *et al.*, 2004), algoritmos de aprendizado de máquinas ou *support vector machines* (SVMs) – (Nagi *et al.*, 2010) –, SVM com lógica *fuzzy* (Nagi *et al.*, 2011), algoritmo de floresta de caminhos ótimos ou *optimum-path forest* (Ramos *et al.*, 2011), método de aprendizado de máquinas extremo ou *extreme learning machine method* (Nizar, Dong e Wang, 2008), algoritmos para formação de agrupamentos ou *clusters* (Angelos *et al.*, 2011).

Neste estudo, implementa-se uma metodologia para detecção de UCs irregulares por meio da aplicação da RNA ARTMAP *Fuzzy*. O conjunto de entrada da rede proposta baseia-se em métricas da estatística descritiva obtidas a partir do histórico de consumo mensal de cada UC. Essas métricas são mais representativas do perfil de consumo de cada UC do que os dados brutos das mesmas em kWh.

A PMC pode ser usada para clarificar algumas vantagens da RNA ARTMAP *Fuzzy*. A PMC é um padrão ou *benchmarking* na área de detecção e classificação de padrões em uma “massa” de dados (Haykin, 2007; Rumerlhart *et al.*, 1986).

Ao treinar a PMC com algoritmo de retropropagação do erro ou *backpropagation*, é esperado que ela responda satisfatoriamente a quaisquer padrões de entrada apresentados no futuro. No entanto, observa-se que o desempenho dessa rede decai gradativamente à medida que o tempo passa, pois os pesos definidos na fase de treinamento não se modificam; portanto, não refletem as novas situações. Para que sempre haja adaptação aos novos padrões de entrada, o algoritmo de aprendizado deve ser *plástico*.

Por outro lado, a PMC pode ser treinada novamente com os novos padrões de entrada. No entanto, isso pode ocasionar a perda de informações aprendidas anteriormente. Para preservar o conhecimento aprendido no passado, o algoritmo de treinamento da rede deve ser *estável*.

Portanto, é preciso um modelo de RNA incremental que não precise reiniciar o treinamento a cada vez que aparecem novos padrões de entrada e que; simultaneamente, preserve o conhecimento adquirido no passado. A família de RNA pertencente à teoria de ressonância adaptativa ou *adaptive resonance theory* (ART) foi desenvolvida para solucionar o dilema entre a estabilidade e a plasticidade.

A *plasticidade* ou aquisição de novos conhecimentos pela rede neural poderia ocasionar a perda de conhecimento previamente adquirido. No entanto, a ARTMAP *Fuzzy* resolve esse dilema e permite a aquisição de novos conhecimentos e; concomitantemente, preserva o conhecimento adquirido anteriormente (Decanini, Tonelli-Neto e Minussi, 2012).

Adicionalmente, a ARTMAP *Fuzzy* tem sido aplicada com sucesso em inúmeros problemas de engenharia envolvendo, por exemplo, a detecção e classificação de faltas de curto-circuito e a previsão de demanda (Santos *et al.*, 2020; Lopes, Minussi e Lotufo, 2005).

Em Faria, Melo e Padilha-Feltrin (2016), as PNTs foram estimadas a nível espaço-temporal por regiões ou áreas de uma cidade. Assim, a abordagem deste estudo de detecção de UCs irregulares pode ser incorporada em trabalhos futuros para detecção de PNTs no interior de áreas vulneráveis da cidade.

Este estudo segue na Seção 2 com uma introdução às RNAs. Na Seção 2.1 aborda-se o algoritmo para treinamento da ARTMAP *Fuzzy*. Na Seção 3 é apresentada a metodologia

proposta: módulo de extração de atributos estatísticos (Seção 3.1), construção do histórico de inspeções (Seção 3.2), sistema classificador (Seção 3.3) e apresentação das métricas para avaliação da metodologia proposta (Seção 3.4). Na Seção 4 são apresentados os resultados após a construção dos cenários de inspeções (Seção 4.1) e a detecção de UCs irregulares (Seção 4.2). Este estudo finaliza com as conclusões na Seção 5.

2. REDES NEURAIS ARTIFICIAIS

Uma das principais características das RNAs é a capacidade de generalização. Trata-se da produção de respostas a padrões de entradas similares àqueles que foram previamente apresentados à rede durante a fase de treinamento (Silva, Spatti e Flauzino, 2014).

As RNAs são unidades de processamento inspiradas no funcionamento do sistema nervoso dos seres vivos, que possuem tendência natural para armazenar conhecimento e torná-los disponíveis para uso. Assemelham-se ao cérebro por dois aspectos. O primeiro refere-se ao conhecimento da rede que é obtido por intermédio do seu ambiente através da aprendizagem. O segundo está associado à força de conexão entre os neurônios que é representada por peso sinápticos. Esses pesos realizam o armazenamento distribuído do conhecimento obtido. As RNAs possuem elevado poder computacional devido à sua estrutura maciçamente paralela e distribuída e por sua capacidade de aprendizado e generalização (Haykin, 2007).

2.1 Rede Neural ARTMAP Fuzzy

A RNA ARTMAP Fuzzy pertence à família ART com treinamento supervisionado e auto-organizável. Sua estrutura é composta por três módulos ART: ART_a , ART_b e ART_{ab} . O módulo ART_a é responsável pelo processamento dos vetores de entrada; enquanto no módulo ART_b , são representadas as respectivas saídas desejadas. Esses dois módulos estão conectados por um terceiro módulo ART_{ab} ou módulo inter-ART (Carpenter *et al.*, 1992).

Essa topologia de rede, possui características de destaque como plasticidade e estabilidade. A *plasticidade* viabiliza a incorporação de novos padrões sem a perda de conhecimentos previamente adquiridos. Enquanto a *estabilidade* está relacionada à preservação do conhecimento via agrupamento de padrões similares na mesma categoria de reconhecimento (Carpenter e Grossberg, 1992).

Quando o módulo ART_a realiza uma estimativa errônea, há uma incompatibilidade entre as categorias ART_a e ART_b , ou seja, a estimativa realizada em ART_a não condiz com a respectiva saída desejada representada no módulo ART_b . Assim, o processo denominado *match-tracking* é ativado por meio da inibição da memória associativa ART_{ab} . Desse modo, o parâmetro de vigilância ρ_a , pertencente ao módulo ART_a sofre alteração, tendo seu valor aumentado a uma quantidade mínima necessária para eliminar totalmente a categoria incompatível. Uma vez que a categoria incompatível foi excluída, o sistema iniciará uma nova busca no respectivo módulo, a fim de encontrar uma categoria compatível com a saída representada em ART_b (Carpenter e Grossberg, 1992).

A Tabela 1 apresenta uma descrição de todos os parâmetros da ARTMAP Fuzzy.

Tabela 1. Parâmetros da ARTMAP Fuzzy.

Parâmetros	Descrição dos Parâmetros
Parâmetro de escolha $\alpha \in \mathbb{R}^+$	Seleção de categorias
Taxa de treinamento $\beta \in [0,1]$	Controla a velocidade de aprendizagem da rede
Parâmetro de vigilância do módulo ART_a $\rho_a \in [0,1]$	Controla a ressonância no módulo ART_a
Parâmetro de vigilância do módulo ART_b $\rho_b \in [0,1]$	Controla a ressonância no módulo ART_b
Parâmetro de vigilância do módulo ART_{ab} $\rho_{ab} \in [0,1]$	Controla a ressonância no módulo ART_{ab}
Constante de valor mínimo $\epsilon \in [0,1]$	Incremento no parâmetro de vigilância do módulo ART_a

2.1.1 Algoritmo para Treinamento da ARTMAP Fuzzy

O algoritmo para treinamento da ARTMAP Fuzzy é apresentado de acordo com Carpenter *et al.* (1992) e Carpenter e Grossberg (1992).

Etapa 1: Dados de Entrada. Nesta etapa, realiza-se a normalização dos dados de entrada via (1):

$$\bar{\mathbf{a}} = \frac{\mathbf{a}}{|\mathbf{a}|} \quad (1)$$

Sendo $|\mathbf{a}| = \sum a_i$, e $\bar{\mathbf{a}}$ representa o vetor de entrada normalizado. Na sequência, é feita a codificação de complemento do respectivo vetor via (2). Esse processo é realizado para preservar a amplitude dos dados. O vetor de entrada \mathbf{I} torna-se $2M$ -dimensional conforme (3).

$$\bar{a}_i^c = 1 - \bar{a}_i \quad (2)$$

$$\mathbf{I} = [\bar{\mathbf{a}} \ \bar{\mathbf{a}}^c] = [\bar{a}_1 \dots \bar{a}_M \ \bar{a}_1^c \dots \bar{a}_M^c] \quad (3)$$

Etapa 2: Vetores de Pesos. Essa topologia de rede possui três conjuntos de vetores de pesos, sendo um para cada módulo ART_a , ART_b e ART_{ab} , conforme apresentado em (4), (5) e (6). Inicialmente todos os elementos dos vetores são iguais e unitários; portanto, não existem categorias ativas.

$$\mathbf{W}_j^a = (W_{j1}^a, W_{j2}^a \dots W_{j,2M_a}^a) \quad (4)$$

$$\mathbf{W}_k^b = (W_{k1}^b, W_{k2}^b \dots W_{k,2M_b}^b) \quad (5)$$

$$\mathbf{W}_{jk}^{ab} = (W_{j1}^{ab}, W_{j2}^{ab} \dots W_{j,2N_b}^{ab}) \quad (6)$$

Etapa 3: Escolha de Categorias. Nesta etapa, realiza-se a escolha de categorias para os módulos ART_a e ART_b , cujas funções de escolha estão em (7) e (8), respectivamente. O operador AND fuzzy “ \wedge ” é definido em (9).

$$T_j^a = \frac{|\mathbf{I}^a \wedge \mathbf{W}_j^a|}{\alpha + |\mathbf{W}_j^a|} \quad (7)$$

$$T_k^b = \frac{|\mathbf{I}^b \wedge \mathbf{W}_k^b|}{\alpha + |\mathbf{W}_k^b|} \quad (8)$$

$$(\mathbf{p} \wedge \mathbf{q})_i \equiv \min(p_i, q_i) \quad (9)$$

A escolha da categoria está associada ao neurônio vencedor. A categoria vencedora J (para ART_a) e K (para ART_b) é obtida de acordo com os critérios apresentados em (10) e (11).

$$J = \arg \{ \max T_j^a \}, \text{ para } j = 1, 2, \dots, N \quad (10)$$

$$K = \arg \{ \max T_k^b \}, \text{ para } k = 1, 2, \dots, N \quad (11)$$

Quando mais de uma categoria for ativada, deve-se escolher aquela que possuir o menor índice j ou k . Após a obtenção do neurônio vencedor (neurônio K), monta-se o vetor de atividade do módulo ART_b $\mathbf{Y}^b = [Y_1^b, \dots, Y_N^b]$, no qual apenas a saída correspondente ao neurônio vencedor K será ativada conforme (12).

$$Y_k^b = \begin{cases} 1, & \text{para } k = K \\ 0, & \text{caso contrário} \end{cases} \quad (12)$$

Etapa 4: Ressonância ou Reset. O processo de ressonância ocorrerá nos módulos ART_a e ART_b se (13) e (14) forem satisfeitas; caso contrário, será efetuado o *reset* nos respectivos módulos. Quando o processo de *reset* é realizado nos módulos ART_a e ART_b , é feita uma busca por uma nova categoria ativa via (10) e (11). Desse modo, os nós J e K , escolhidos anteriormente, serão excluídos do novo processo de busca; assim, os valores de T_j^a e T_k^b são fixados em zero – $T_j^a = 0$ e $T_k^b = 0$. Esse procedimento repetir-se-á até que seja encontrada uma categoria em cada módulo que satisfaça (13) e (14) para conclusão do processo de ressonância.

$$\frac{|I^a \wedge W_j^a|}{|I^a|} \geq \rho_a \quad (13)$$

$$\frac{|I^b \wedge W_k^b|}{|I^b|} \geq \rho_b \quad (14)$$

Etapa 5: Match-Tracking. Nesta etapa é avaliado se o neurônio ativado no módulo ART_a possui compatibilidade com a saída desejada no módulo ART_b . Em (15) é avaliado se o critério de vigilância é satisfeito.

Se (15) for satisfeita, então o processo de ressonância no módulo ART_a será confirmado e os pesos serão atualizados; caso contrário, será introduzido um pequeno acréscimo em ρ_a via (16), T_j^a recebe valor nulo e inicia-se um novo processo de busca por uma nova categoria no módulo ART_a por meio de (10). Esse processo é realizado até que um nó escolhido em ART_a satisfaça o critério de combinação em (15).

$$\frac{|Y^b \wedge W_j^{ab}|}{|Y^b|} \geq \rho_{ab} \quad (15)$$

$$\rho_a = \frac{|I^a \wedge W_j^a|}{|I^a|} + \epsilon \quad (16)$$

O vetor de atividade no módulo ART_a $\mathbf{Y}^a = [Y_1^a, \dots, Y_N^a]$ é obtido via (17) e é construído após a confirmação da categoria vencedora via ação do dispositivo *match-tracking*.

$$Y_j^a = \begin{cases} 1, & \text{para } j = J \\ 0, & \text{caso contrário} \end{cases} \quad (17)$$

Etapa 6: Aprendizado. Se o critério de vigilância no módulo ART_{ab} (15) for cumprido, os pesos dos módulos serão atualizados de acordo com (18), (19) e (20) e o processo de aprendizado é realizado.

$$W_j^{a(novo)} = \beta(I^a \wedge W_j^{(anterior)}) + (1 - \beta)W_j^{(anterior)} \quad (18)$$

$$W_k^{b(novo)} = \beta(I^b \wedge W_k^{(anterior)}) + (1 - \beta)W_k^{(anterior)} \quad (19)$$

$$W_{jk}^{ab(novo)} = \begin{cases} 1, & \text{se } j = J \text{ e } k = K \\ 0, & \text{caso contrário} \end{cases} \quad (20)$$

3. DETECÇÃO DE PERDAS NÃO TÉCNICAS

Nesta seção são apresentados os detalhes da metodologia para identificação PNTs a partir do histórico de consumo mensal em kWh via RNA ARTMAP *Fuzzy*.

3.1 Atributos Estatísticos Baseados em Regimes

Nesta seção mostra-se como obter 16 (dezesseis) atributos estatísticos a partir do histórico de consumo mensal em kWh. Esses dados são mais representativos do perfil de consumo de cada UC, com menor dimensão do que os dados originais e sem redundâncias. Dados estatísticos são mais intuitivos e; portanto, de mais fácil compreensão por um especialista em perdas. Esses atributos serão os dados de entrada da ARTMAP *Fuzzy* em lugar dos dados brutos de consumo de cada UC.

Alguns trabalhos em PNTs extraem coeficientes *Wavelets* e de Fourier da série de consumo mensal (Cometti e Varejão, 2005). Em geral, esses coeficientes são as entradas de técnicas de *soft computing* como: mineração de dados, RNAs, lógica *fuzzy*, algoritmo de aprendizado de máquinas ou SVMs e outros (Cabral *et al.*, 2004; Messinis e Hatziargyriou, 2018; Nagi *et al.*, 2010; Nagi *et al.*, 2011; Nizar, Dong e Wang, 2008). Ocorre que os coeficientes *Wavelets* e de Fourier são capazes de detectar mínimas variações na curva de consumo mensal; portanto, são mais suscetíveis para captar ruídos inerentes à curva de consumo mensal.

Por outro lado, os atributos estatísticos baseados em regimes fornecem uma visão geral da tendência de consumo da UC: queda, elevação ou estagnação dentro de uma faixa central. Assim, são menos sensíveis às pequenas oscilações na curva de consumo mensal que reduzem o desempenho dos sistemas classificadores.

A Tabela 2 contém a descrição dos dezesseis atributos estatísticos baseados no número de regimes do histórico de consumo mensal em kWh (Faria, Padilha-Feltrin e Minussi, 2012; Ferreira, 2008).

Todas as métricas da Tabela 2 são baseadas em regimes; portanto, convém mostrar a obtenção de regimes a partir de uma série de consumo mensal. Nesse sentido, a Tabela 3 contém um exemplo ilustrativo do consumo anual de uma UC hipotética. Observa-se que essa série possui três regimes. Os primeiros 4 meses estão no 1º regime e têm média (ou patamar) de 395 kWh. Todos os meses pertencentes a esse primeiro regime variam, no máximo, 25% em relação à média do regime. Essas observações aplicam-se aos demais regimes.

Tabela 2. Atributos estatísticos obtidos a partir da curva de consumo mensal em kWh.

Parâmetros	Descrições
<i>Regimes</i>	N.º de regimes ou patamares do histórico de consumo mensal.
<i>CV</i>	Coefficiente de Variação. Representa a variabilidade da série de consumo mensal em relação à sua média. $CV = \frac{\text{Desvio Padrão}}{\text{Média}}$
<i>PQR</i>	Percentual de queda de regime (<i>PQR</i>) e de aumento de regime (<i>PAR</i>). Na mudança de regime é verificado se o regime subsequente será de queda ou de aumento.
<i>PAR</i>	
<i>NQR</i>	N.º de queda de regime (<i>NQR</i>) e número de aumento de regime (<i>NAR</i>). Na mudança de regime é verificado se o regime subsequente será de queda ou de aumento.
<i>NAR</i>	
<i>PMRI</i>	Percentagem em relação ao número total de meses em que o consumo esteve no regime inicial (<i>PMRI</i>), em regimes de queda (<i>PMRQ</i>) e em regimes de aumento (<i>PMRA</i>).
<i>PMRQ</i>	
<i>PMRA</i>	
<i>NZ</i>	N.º de zeros. Consumo inferior a 10% da média da série.
<i>NRFM</i>	N.º de Regimes na Faixa Média.
<i>NRAbFM</i>	N.º de Regimes Abaixo da Faixa Média.
<i>NRAcFM</i>	N.º de Regimes Acima da Faixa Média.
<i>NRFRI</i>	N.º de Regimes na Faixa do Regime Inicial.
<i>NRAbRI</i>	N.º de Regimes Abaixo do Regime Inicial.
<i>NRACRI</i>	N.º de Regimes Acima do Regime Inicial.

Os dados da Tabela 3 estão representados graficamente na Fig. 2 onde a curva azul e a curva vermelha representam, respectivamente, o histórico de consumo mensal em kWh e os três regimes ou patamares da mesma ao longo de 12 meses.

3.2 Construção do Histórico de Inspeções de PNTs

Vital para o treinamento (e aprendizado) do sistema classificador ARTMAP *Fuzzy* é o histórico de inspeções que contém as amostras de UCs regulares e irregulares.

Tabela 3. Exemplo ilustrativo para obtenção de regimes a partir da série de consumo mensal em kWh.

Série de Consumo Mensal [kWh]	Regimes	Patamares [kWh]
336	1º	395
409	1º	
433	1º	
402	1º	
512	2º	455
398	2º	
306	3º	301
312	3º	
290	3º	
271	3º	
317	3º	
310	3º	

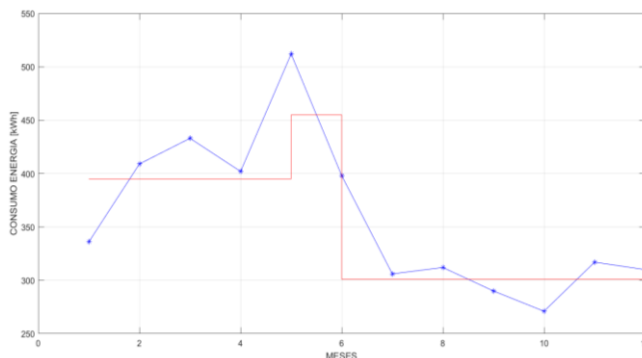


Fig. 2. Curva de consumo mensal hipotética em kWh e os regimes (ou patamares) da mesma.

Como os autores não têm acesso aos dados de consumo das UCs inspecionadas, será construído um histórico de inspeções a partir de um histórico de consumo real em kWh.

O histórico de inspeções será construído via algoritmo para formações de agrupamentos ou *clusters* de amostras semelhantes. Portanto, será formado um *cluster* majoritário com UCs regulares e outro *cluster* minoritário com UCs irregulares com consumo atípico. Observa-se que, a fim de avaliar a eficiência do sistema classificador proposto, serão inseridas amostras de UCs com consumo regular no *cluster* minoritário de UCs irregulares e vice-versa.

Neste estudo, foi aplicada a RNA auto-organizável de Kohonen. Ela é adequada para construção de *clusters* com características semelhantes. Outros algoritmos para construção de *clusters* poderiam ter sido aplicados como o *k-means*, por exemplo.

A rede de Kohonen é um tipo de RNA com aprendizado não supervisionado, competitivo e que identificam padrões em vetores de dados multivariados. Essa rede é capaz de captar o padrão dos dados de entrada via vetores de pesos neurais.

Detalhes adicionais a respeito da rede neural auto-organizável de Kohonen foge ao escopo deste estudo. Para o leitor interessado recomenda-se o capítulo 8 de Silva, Spatti e Flauzino (2014) e o capítulo 9 de Haykin (2007).

3.3 Classificação de UCs da Série de Consumo Mensal

De posse do histórico de inspeções construído conforme mostrado na seção anterior, pode-se aplicá-lo para aprendizado e teste do sistema classificador proposto neste estudo via ARTMAP *Fuzzy*.

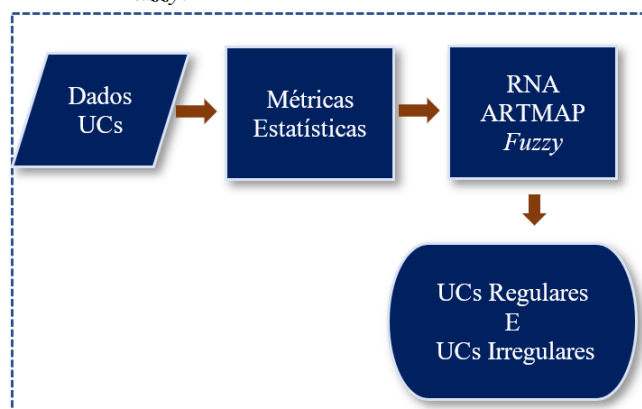


Fig. 3. Fluxograma da metodologia proposta para detecção de PNTs.

Essa rede é capaz de identificar padrões ocultos via adaptações de seus pesos neurais. Assim pode-se identificar UCs possivelmente irregulares e aumentar a taxa de sucesso das inspeções em campo.

A Fig. 3 apresenta um fluxograma com a visão geral da metodologia proposta neste estudo para detecção de PNTs no SDEE.

3.4 Validação: Métricas para Avaliar o Desempenho do Sistema Classificador

Neste estudo, desenvolve-se uma metodologia onde se aplica um sistema classificador baseado na rede ARTMAP Fuzzy para detecção de PNTs.

A métrica usualmente aplicada para avaliar o desempenho de sistemas classificadores é a *Taxa de Sucesso (TS)* conforme (21). No entanto, apenas essa métrica é insuficiente para estudos com PNTs, pois as duas classes de UCs de regulares e UCs irregulares são muito desiguais. Em média, nas inspeções em campo, para cada UC irregular, há nove UCs regulares (Cometti e Varejão, 2005).

Nesse sentido, como ilustração, seja um classificador hipotético com desempenho insatisfatório. Se ele classificar todas as UCs como sendo regulares, então ele apresentará uma alta *TS* sem acertar um único caso de fraude!

Portanto, são avaliadas outras métricas em conjunto com a *TS*. A saber: a confiabilidade negativa (*CN*) em (22) e a especificidade (*ESP*) em (23) que darão suporte para melhor avaliação do sistema classificador proposto neste estudo.

Essas métricas foram obtidas a partir da matriz de confusão apresentada na Tabela 4. Essa matriz foi adaptada para o problema de PNTs e contém duas classes: UCs regulares ou normais (N) e UCs irregulares ou com fraude (F).

Tabela 4. Matriz de confusão.

Matriz de Confusão		Classe Predita	
		N	F
Classe Real	N	Q_{NN}	Q_{NF}
	F	Q_{FN}	Q_{FF}

Sendo:

Q_{NN} : número de UCs regulares que foram corretamente classificadas – verdadeiro positivo.

Q_{NF} : número de UCs regulares, mas que foram incorretamente classificados como UCs irregulares – falso negativo.

Q_{FN} : número de UCs irregulares, mas que foram incorretamente classificadas como regulares – falso positivo.

Q_{FF} : número de UCs irregulares e que foram corretamente classificadas – verdadeiro negativo.

A *CN* em (22) é a razão entre o número total de UCs irregulares corretamente classificadas e o número de UCs (corretamente e incorretamente) classificadas como irregulares.

Por outro lado, a *ESP* em (23) é a razão entre o número de UCs irregulares corretamente classificadas e o número total de UCs irregulares contidas na base de dados.

$$TS = \frac{Q_{NN} + Q_{FF}}{Q_{NN} + Q_{FF} + Q_{NF} + Q_{FN}} \quad (21)$$

$$CN = \frac{Q_{FF}}{Q_{FF} + Q_{NF}} \quad (22)$$

$$ESP = \frac{Q_{FF}}{Q_{FF} + Q_{FN}} \quad (23)$$

4. RESULTADOS

Neste estudo é proposta uma metodologia para detecção de UCs irregulares (possivelmente com PNTs) localizadas em uma rede de distribuição real localizada no interior paulista. Consideram-se dados reais de UCs ao longo de 36 meses.

O código para detecção de PNTs foi desenvolvido no ambiente MATLAB® (MATLAB, 2010). No entanto, o mesmo pode ser migrado para quaisquer linguagens de alto nível como: C++, Fortran ou Python.

Todas as simulações apresentadas foram executadas em um computador pessoal com processador Intel Core i7; 1,8–2,3 GHz e 16 GB de memória RAM.

4.1 Construção de Cenários de PNTs

Conforme abordado na Seção 3.2, neste estudo aplica-se a rede auto-organizável de Kohonen para construção de agrupamentos ou *clusters* de UCs regulares e irregulares a partir do histórico de consumo real das UCs em kWh. Os parâmetros para aplicação da rede de Kohonen estão apresentados na Tabela 5.

A Tabela 6 contém informações adicionais a respeito do histórico de inspeções construído com 1.000 UCs residenciais constituído por 900 UCs regulares (*cluster 1*) e 100 UCs irregulares (*cluster 2*).

A proporção de UCs em cada *cluster* está de acordo com a literatura especializada em PNTs, pois, conforme Cometti e Varejão (2005), em média a cada 10 UCs inspecionadas, apenas uma estará efetivamente irregular.

A Tabela 6 contém a média aritmética dos parâmetros estatísticos para todas as UCs contidas em cada *cluster* de UCs regulares e irregulares. Apenas a média aritmética não é capaz de revelar o perfil completo de cada *cluster*. No entanto, pode-se aplicá-la para concluir que o *cluster 1* apresenta UCs com consumo “bem-comportado” e regular do que o *cluster 2*.

Tabela 5. Parâmetros da rede auto-organizável de Kohonen.

Parâmetros	Descrição
Taxa de Treinamento α	0,5
Constante de Convergência ϵ	10^{-2}
Número de Épocas para Treinamento da rede	10^3
Treinamento	Aprendizado Competitivo
Tipo de Treinamento	Aprendizado Não Supervisionado
Critério de Convergência	Varição da magnitude dos vetores de pesos da RNA

Tabela 6. Média dos atributos estatísticos para cada agrupamento (clusters) de UCs.

Parâmetros	Cluster 1 (UCs Regulares)	Cluster 2 (UCs Irregulares)
<i>N.º UCs</i>	900	100
<i>Regimes</i>	7,96	4,8
<i>CV</i>	0,33	3,09
<i>PQR</i>	0,44	0,09
<i>PAR</i>	0,36	0,91
<i>NQR</i>	3,11	1,62
<i>NAR</i>	3,85	2,18
<i>PMRI</i>	0,27	0,50
<i>PMRQ</i>	0,26	0,33
<i>PMRA</i>	0,46	0,17
<i>NZ</i>	1,72	27,39
<i>NRFM</i>	2,34	1,62
<i>NRAbFM</i>	3,31	0,79
<i>NRAcFM</i>	2,30	2,39
<i>NRFRI</i>	2,03	1,55
<i>NRAbRI</i>	3,02	0,61
<i>NRAcRI</i>	2,90	2,64

Essa conclusão fica evidente a partir dos valores dos coeficientes de variação (*CV*) e números de zeros (*NZ*) do *cluster 1* que são bem inferiores àqueles aferidos para o *cluster 2*. Portanto, a partir dessa análise, atribuímos ao *cluster 1* e ao *cluster 2* as UCs regulares e irregulares, respectivamente.

4.2 Detecção de Perdas Não Técnicas

No processo para detecção de PNTs, as métricas estatísticas da Tabela 2 constituem as entradas do sistema classificador formado pela ARTMAP *Fuzzy*. Essa rede fornece em seu vetor de saída as respectivas UCs regulares e UCs irregulares. A Tabela 7 contém todos os parâmetros adotados para o sistema classificador ARTMAP *fuzzy*.

O conjunto de dados é conformado por 1.000 UCs. A partir dessa base de dados inicial formaram-se outros dois conjuntos: *conjunto de treinamento* com 700 UCs (70% do número total de UCs) e *conjunto de teste* com 300 UCs (30% do total de UCs). Observa-se que nos dois conjuntos de treinamento e de teste há uma mesma proporção de UCs regulares e de UCs irregulares de 10:1.

A Tabela 8 contém os resultados numéricos para a matriz de confusão. Esses valores foram obtidos a partir das 300 UCs do conjunto de teste a fim de avaliar o desempenho do classificador proposto.

Tabela 7. Parâmetros da RNA ARTMAP Fuzzy.

Parâmetros	Valores
Parâmetros de Escolha α	0,1
Taxa de Treinamento β	0,5
Parâmetro de Vigilância do Módulo $ART_a \rho_a$	0,8
Parâmetro de Vigilância do Módulo $ART_b \rho_b$	0,9
Parâmetro de Vigilância do Módulo $ART_{ab} \rho_{ab}$	0,8
Constante ϵ	10^{-5}

Tabela 8. Matriz de confusão com valores numéricos.

Matriz de Confusão		Classe Predita	
		N	F
Classe Real	N	$Q_{NN} = 268$	$Q_{NF} = 2$
	F	$Q_{FN} = 1$	$Q_{FF} = 29$

A partir de (21)–(23) tem-se os valores obtidos pelo sistema classificador: $TS = 99\%$, $CN = 93,6\%$ e $ESP = 96,7\%$, respectivamente.

A especificidade (*ESP*) expressa a cobertura do sistema classificador, isto é, o percentual de UCs irregulares que o sistema consegue identificar.

A confiabilidade negativa (*CN*) é um indicativo da precisão das inspeções a serem realizadas, isto é, o percentual de êxitos na identificação de UCs irregulares em relação ao total de inspeções realizadas.

Um classificador com alta *ESP* e baixa *CN* possui boa cobertura, mas pouco êxito nas inspeções. Por outro lado, um classificador com alta *CN* e baixa *ESP* possui boa taxa de sucesso nas inspeções, mas não identifica uma parcela significativa das UCs irregulares existente no SDEE em estudo.

Portanto, o classificador proposto neste estudo conformado pela ARTMAP *Fuzzy* obteve valores elevados para todas as métricas avaliadas: *TS*, *CN* e *ESP*. Logo, o mesmo possui boa cobertura na identificação de UCs irregulares e promove inspeções em campo com mais sucesso, ao minimizar visitas desnecessárias das equipes em campo às UCs regulares ou normais.

6. CONCLUSÃO

Neste estudo foi apresentada uma metodologia para combate às perdas não técnicas (PNTs) no sistema de distribuição de energia elétrica (SDEE) via detecção de unidades consumidoras (UCs) irregulares. A metodologia proposta tem dois módulos: (i) extração de atributos estatísticos baseados em regimes e (ii) módulo de classificação baseado na rede neural artificial (RNA) ARTMAP *Fuzzy*.

A metodologia mostrou-se viável para identificação de UCs irregulares, pois apresentou valores elevados para todas as métricas aplicadas ao conjunto de testes: confiabilidade negativa ($CN = 93,6\%$), especificidade ($ESP = 96,7\%$) e taxa de sucesso ($TS = 99\%$).

Portanto, o sistema classificador proposto apresentou valor elevado para *ESP* o que indica boa cobertura ao identificar uma parcela relevante das UCs irregulares no SDEE em estudo. Além disso, o valor elevado para *CN* é relevante, pois minora o problema de as equipes de inspeções visitarem inadvertidamente UCs regulares ou normais.

Como trabalhos futuros sugere-se o uso de bases de dados com uma quantidade maior de UCs e pertencentes a outras classes de consumidores da rede de distribuição como UCs da classe comercial, por exemplo.

Além disso, trabalhos com detecção aprimorada de UCs irregulares podem ser incorporados em outros estudos que detectam áreas endêmicas e mais vulneráveis às PNTs.

AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pela concessão da bolsa de Iniciação Científica (Processo: 21/03103-1) e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) – Código de financiamento 001 – pela concessão de uma bolsa de doutorado.

REFERÊNCIAS

- Agência Nacional de Energia Elétrica (ANEEL) (2020). Perdas de Energia Elétrica na Distribuição. Disponível em: <<https://www.aneel.gov.br/>>.
- Angelos, E. W. S., Saavedra, O. R., Cortés, O. A. C., Souza, A. N. (2011). Detection and Identification of Abnormalities in Customer Consumptions in Power Distribution Systems. *IEEE Trans. on Power Delivery*, v. 26.
- Associação Brasileira de Distribuidores de Energia Elétrica (ABRADEE) (2018). *Furto e Fraude de Energia*. Disponível em: <<https://www.abradee.org.br/>>.
- Cabral, J. E., Pinto, J. O., Gontijo, E. M., Filho, J. R. (2004). Fraud detection in electrical energy consumers using rough sets. *IEEE International Conference on Systems, Man and Cybernetics*.
- Carpenter, G.A.; Grossberg, S.; Markuzon, N.; Reynolds, J.H.; Rosen, D.B. (1992). *Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps*. *IEEE Trans. on Neural Networks*, v. 3, p. 698-713.
- Carpenter, G.A., Grossberg, S. (1992). A self-organizing neural network for supervised learning, recognition and prediction. *IEEE Communications Magazine*, v. 30, n. 9, p. 38-49.
- Cometti, E. S., Varejão, F. M. (2005). Melhoramentos na identificação de perdas comerciais através da análise computacional inteligente do perfil de consumo e dos dados cadastrais de consumidores. Vitória-ES: Relatório técnico final do projeto de P&D.
- Decanini, J.G.M.S., Tonelli-Neto, M.S., Minussi, C.R. (2012). Robust fault diagnosis in power distribution systems based on *fuzzy ARTMAP* neural network-aided evidence theory. *IET Generation, Transmission & Distribution*, v. 6, p. 1112-1120.
- Haykin, S. (2007). *Redes neurais: princípios e prática*. Bookman Editora.
- Faria, L.T., Melo, J.D., Padilha-Feltrin, A. (2016). Spatial-Temporal Estimation for Non-Technical Losses. *IEEE Trans. Power Delivery*, v. 31, n. 1, p. 362-369.
- Faria, L.T., Padilha-Feltrin, A., Minussi, C.R. (2012). Sistema inteligente híbrido intercomunicativo para detecção de perdas comerciais em sistemas de distribuição de energia elétrica. In IV Simpósio Brasileiro de Sistemas Elétricos – IV SBSE, Goiânia-GO.
- Ferreira, H.M (2008). Uso de ferramentas de aprendizado de máquina para prospecção de perdas comerciais em distribuição de energia elétrica. 2008. 89 f. Dissertação de Mestrado. Faculdade de Engenharia Elétrica e de Computação, Unicamp, Campinas, São Paulo.
- Lopes, M.L.M., Minussi, C.R., Lotufo, A.D.P. (2005). Electric load forecasting using a fuzzy ART&ARTMAP neural network. *Applied soft computing*, v. 5, p. 235-244.
- MATLAB (2010). Version 7.10.0 (R2010a), Natick, Massachusetts: The MathWorks Inc.
- Messinis, G.M., Hatziargyriou, N.D. (2018). Review of non-technical loss detection methods. *Electrical Power Systems Research*, p. 250-266.
- Nagi, J., Yap, K.S., Tiong, S.K., Ahmed, S.K., Mohamad, M. (2010). Nontechnical Loss Detection for Metered Customers in Power Utility Using Support Vector Machines. *IEEE Trans. Power Delivery*, v. 25, p. 1162-1171.
- Nagi, J., Yap, K.S., Tiong, S.K., Ahmed, S.K., Nagi, F. (2011). Improving SVM-Based Nontechnical Loss Detection in Power Utility Using the *Fuzzy Inference System*. *IEEE Trans. Power Delivery*, v. 26, p. 1284-1285.
- Nizar, A. H., Dong, Z. Y., Wang, Y. Power Utility Nontechnical Loss Analysis with Extreme Learning Machine Method. *IEEE Trans. Power Systems*, v. 23, 2008.
- Ramos, C. C. O., Sousa, A. N., Papa, J. P., Falcão, A. X. A (2011). New Approach for Nontechnical Losses Detection Based on Optimum-Path Forest. *IEEE Trans. Power Systems*, v. 26.
- Rumerlhart, D.E., Hinton, G.E., Williams, R.J. (1986). *Learning internal representations by error propagation*. In: *Parallel Distributed Processing*, v. 1, chapter 8. MIT Press, Cambridge, Massachusetts, USA.
- Santos, A.S., Bernardes, H.R., Faria, L.T., Lopes, M.L., Minussi, C.R. (2020). Detecção e classificação de faltas de curto-circuito em sistemas de distribuição de energia elétrica com inserção de geração distribuída. In XXIII Congresso Brasileiro de Automática – XXIII CBA, Congresso Virtual.
- Silva, I.N., Spatti, D.H., Flauzino, R.A. (2014). *Redes Neurais Artificiais: para engenharias e ciências aplicadas*. São Paulo: Ed. Artiliber, 2ª ed.