

# Identification of Similar Loads for Electric Power Management in Smart Grid

R. Barcelos, W. Cardoso, L. Rigo Jr., and G. Diniz

**Abstract**—This paper deals with the identification of highly similar loads in Smart Grids. The goal is to apply load characterization techniques to define a unique load signature. In this work is used a platform composed of four fluorescent lamps with the same technical data and from the same manufacturer. Fluorescent lamps can be driven individually or together allowing for 16 different operating configurations, this means, from no one lamp turned on to all the lamps turned on. For each configuration, 150 voltage cycles and 150 cycles of total circuit current are collected at a frequency of 25 kHz. Then, two forms of load signatures are tested, one where the signature is composed of 14 characteristics and another one, based on the entropy of Shannon and Renyi, where the signature can have 5, 15, 45 or 85 characteristics. After defining the signature, it is necessary to go through the classification system to find the lowest error rate of the identification system. In this work, five classifiers were tested. The lowest error rate found was 25.63% and is close to the literature when the scenario is composed of different loads, thus proving the efficiency of the presented method.

**Index Terms**—Load Identification, Load Signature, Smart Grid.

## I. INTRODUÇÃO

O Consumo mundial de energia elétrica, em 2006, era de 15.794 TWh, segundo dados da Agência Internacional de Energia (*International Energy Agency* – IEA). Em 2016 esse consumo aumentou para 20.863 TWh, representando um acréscimo de 32,09% no consumo de energia elétrica em 10 anos [1]. O consumo de energia elétrica está diretamente associado à evolução tecnológica, ao crescimento econômico de um país, ao conforto e à dignidade humana. Por isso, o acesso a essa forma de energia é tido como essencial e um direito universal [2]. Considerando o caráter indispensável desse recurso, o aumento do seu consumo nos últimos anos e a finitude das fontes de energia, é necessário pensar em formas sustentáveis de produzir e consumir essa energia.

A aplicação da tecnologia da informação em redes de energia elétrica configurou o que é conhecido como rede inteligente ou *Smart Grid*. Essas redes possibilitam que as informações obtidas sejam utilizadas para o gerenciamento ótimo da energia elétrica, tornando seu uso mais eficiente. Estudos feitos com residências comprovaram que o gerenciamento de cargas elétricas pode economizar de 10% a 15% da energia elétrica consumida [3]. Entretanto, para gerenciar as

cargas de uma *Smart Grid* é necessário identificar essas cargas no sistema.

Uma das maneiras de identificar cargas é instalar um sensor em cada uma delas (*Intrusive Load Monitoring* – ILM). Assim, é possível saber, a qualquer instante, qual carga está consumindo energia elétrica e quanto está consumindo. Inicialmente, esse método requer gastos com a instalação de sensores, fios e manutenção [4]. Ademais, cada novo equipamento introduzido na instalação elétrica requer mais infraestrutura de sensoriamento, o que pode não ser uma solução prática. Com o advento da “internet das coisas” (*Internet of Things* – IoT) é possível resolver o problema de sensoriamento presente na metodologia ILM, uma vez que boa parte dos objetos presentes no nosso dia a dia possui tecnologias embutidas como sensores e comunicação *wireless* [5]. Com isso, a comunicação pode ser explorada para interconectar e trocar dados entre os dispositivos e sensores, diminuindo os custos de instalação e manutenção.

Além da metodologia ILM, outra forma de monitoramento tem sido desenvolvida, onde apenas um único conjunto de sensores é suficiente para coletar dados de todos os equipamentos ligados em uma instalação elétrica. Essa metodologia foi proposta no início dos anos 90 por Hart e é conhecida como *Non-Intrusive Load Monitoring* (NILM) [6].

O monitoramento não-intrusivo consiste em coletar os dados do sinal de entrada do sistema e desagregar esse sinal total, de modo a identificar quais cargas, pertencentes a uma base conhecida, estão ligadas e quanto de energia elétrica estão consumindo [4]. Portanto, é necessário caracterizar cada carga, atribuindo a cada equipamento uma “assinatura elétrica” específica.

A metodologia NILM necessita, portanto, de um sistema de identificação, que passa por duas etapas: a primeira que define e extrai as características dos sinais elétricos adquiridos pelo sistema NILM os quais compõem a “assinatura”; a segunda etapa que consiste em usar tais “assinaturas” como referência ao comparar com os dados obtidos pelo sistema NILM em tempo real.

Uma análise da primeira etapa mostra que, na busca por essa “assinatura”, diversas técnicas têm sido testadas como, por exemplo, *Vector Projection Classification* (VPC) [7], transformada de Fourier e transformada *Wavelet*, entre outras, não havendo ainda consenso sobre um conjunto de características ótimo [8], [9]. As “assinaturas elétricas” podem ser obtidas aplicando técnicas de extração de características no estado transitório ou no estado estacionário do sinal analisado, dividindo o estudo em duas vertentes.

As “assinaturas elétricas” são extraídas e armazenadas em

R. K. S. Barcelos was with the Federal Institute of Brasília, Brazil, e-mail: 3020422@etfbsb.edu.br.

W. C. Celeste and L. O. Rigo Jr. were with the Department of Computer Science and Electronics of Federal University of Espírito Santo, Brazil.

G. L. D. Chaves was with the Department of Engineering and Technology of Federal University of Espírito Santo, Brazil.

um banco de dados. Quando o sistema está em plena operação, o sinal elétrico total do circuito é coletado por sensores e as características desse sinal são extraídas e comparadas com as assinaturas presentes no banco de dados. Portanto, a segunda etapa consiste em comparar a “assinatura” do sinal total coletado com as “assinaturas” individuais, e suas combinações, das cargas presentes na rede, visando definir, pela análise do sinal total, quais cargas estão ligadas. Técnicas de classificação como *K-Means*, raciocínio baseado em casos (RBC) e redes neurais são utilizadas nessa segunda etapa do sistema de identificação de cargas elétricas. Essas e outras técnicas também vêm sendo estudadas para melhorar a acurácia dos sistemas de identificação via NILM.

O sistema de identificação como um todo passa por essas duas etapas. A primeira é definida aqui como etapa de caracterização e a segunda como etapa de classificação. As duas etapas são complementares, entretanto o trabalho que é apresentado neste artigo explora essencialmente o sinal em regime estacionário e a primeira etapa do processo de identificação, isto é, a caracterização através da definição da “assinatura” da carga.

Normalmente os testes feitos para apurar a acurácia do sistema de identificação consistem em um cenário de cargas com comportamentos e dados técnicos distintos entre si. Com isso, o estado da arte tem deixado em aberto algumas questões específicas como, por exemplo, o problema de identificar cargas com alto grau de similaridade. Em uma indústria, por exemplo, pode ser comum o uso de motores do mesmo tipo e potência. Outro exemplo, são os alimentadores de uma rede de distribuição onde existem diversas residências conectadas, elas podem ser representadas como cargas elétricas altamente similares entre si. Em uma residência, por exemplo, podem existir dois chuveiros com a mesma potência, ou outros equipamentos em duplicidade, como televisões, configurando um cenário com cargas altamente similares. Dessa forma, a presença de cargas altamente similares não é incomum e precisa ser explorada para se ter um controle mais apurado sobre as cargas do sistema analisado. Além disso, é uma prática comum, em grande parte dos trabalhos afins encontrados na literatura, que as cargas sejam acionadas individualmente, isto é, uma de cada vez, facilitando o processo de identificação, mas se afastando do comportamento real normalmente encontrado, onde os sinais de potência são demandados por múltiplos equipamentos.

O trabalho que é apresentado neste artigo se diferencia dos demais nesses dois pontos: o primeiro é que aqui o cenário é composto por cargas com os mesmos dados técnicos e mesmo comportamento, sendo, portanto, definidas como altamente similares; a segunda grande diferença é que a identificação é feita sobre sinais provenientes de cargas acionadas tanto individualmente quanto de forma múltipla, representando assim um comportamento mais próximo do mundo real. Esses dois princípios negligenciados nos diversos trabalhos afins encontrados na literatura elevam, em muito, o grau de dificuldade do problema de identificação, porém agregam componentes que aproximam o problema à realidade observada e contribuem para o avanço do estado da arte na identificação de cargas elétricas em *Smart Grids*. Portanto, o artigo contribui ao tratar

de cargas altamente similares, que é um problema ainda em aberto em *Smart Grid*, o qual é conhecido, mas ainda pouco explorado.

A literatura mostra que sistemas de identificação envolvendo cargas pouco similares têm apresentado bons resultados. Porém, a simples presença de uma única carga altamente similar pode comprometer tais sistemas. Isso se observa a partir da pesquisa bibliométrica detalhada na seção II e que mostra a inexistência de artigos escritos em língua inglesa que tratam de cargas com um alto grau de similaridade, sendo encontrado apenas o trabalho apresentado em [10] escrito em língua portuguesa, o qual se utiliza de lâmpadas com especificações técnicas idênticas em procedimentos experimentais de identificação inteligente. Ainda em [10] o autor varia os métodos de classificação aplicados a um protótipo com quatro lâmpadas fluorescentes e mensura a acurácia do sistema de identificação utilizando 14 características como “assinatura” das cargas (no caso, as lâmpadas fluorescentes com as mesmas especificações técnicas). No trabalho desenvolvido neste artigo, utiliza-se o banco de dados obtidos em [10], de onde são extraídas novas características dos equipamentos, as quais são submetidas aos mesmos classificadores usados em [10], porém com as devidas adaptações.

Logo, é importante ressaltar que o uso de 4 lâmpadas fluorescentes compactas altamente similares se mostrou conveniente por serem comuns em instalações elétricas e serem de baixo custo. Entretanto, não há impedimentos para que as cargas utilizadas fossem outras como, por exemplo, ventiladores, máquinas de lavar roupas ou computadores, desde que apresentassem alto grau de similaridade.

Para a etapa de caracterização são comparadas duas técnicas baseadas em ferramentas consolidadas na literatura. Para a segunda etapa são utilizados cinco classificadores distintos (os mesmos utilizados em [10], porém com pequenas adaptações quando necessário). As técnicas de caracterização, os classificadores e a plataforma experimental utilizada estão descritos na seção III, enquanto que a análise e os resultados obtidos são mostrados na seção IV. A seção V traz as considerações finais.

## II. PESQUISA BIBLIOMÉTRICA

A busca por artigos que tratem de cargas “iguais” ou semelhantes é feita no banco do *Institute of Electrical and Electronics Engineers* (IEEE) e no portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) através de acesso provido pela Universidade Federal do Espírito Santo (UFES) até abril de 2019. Ambos são bancos renomados de publicações científicas, sendo que o portal da CAPES permite a busca em diversos locais como, por exemplo, o *Web of Science*, o próprio IEEE, o Springer e a Elsevier. Nos referidos portais, foi utilizada a busca em todo o texto (*full text*) estando no modo avançado e usando-se quatro conjuntos de palavras-chave, a saber: (i) “*load signature\**” and “*similar load\**”; (ii) “*load signature\**” and “*equal load\**”; (iii) “*load identification\**” and “*similar load\**”; e, (iv) “*load identification\**” and “*equal load\**”. O asterisco permite encontrar tanto *load* quanto *loads*. O assunto assinatura de

carga - “load signature” - é extenso, mas com o filtro adicional de carga similar - “similar load” - é possível identificar apenas os artigos que, em algum momento, comentaram sobre cargas similares ou idênticas. Os outros campos da busca são deixados o mais abrangente possível. O campo de intervalo de tempo, por exemplo, é deixado desde a data mais antiga até a data mais recente disponível. Ressalta-se que é pesquisado um conjunto de palavras-chave por vez.

A partir da busca conforme descrito acima, chegou-se a 30 artigos. Desses, 12 artigos tratam de assuntos como redes neurais, padrão de consumo pelo lado da demanda ou reservas de energia, não dando enfoque ao problema de identificação inteligente de cargas elétricas. Já os dezoito artigos restantes, a saber [11]–[28], concentram o estado da arte referente ao problema de identificação inteligente de cargas em instalações elétricas. Em tais trabalhos, os autores tratam de analisar componentes harmônicas, ou de fazer comparativos entre os diversos métodos de caracterização de cargas existentes ou ainda de buscar formas de diminuir o banco de dados de assinaturas visando acelerar o processo de identificação. Entretanto, em nenhum desses trabalhos são analisadas cargas com um grau de similaridade tão alto como o que é abordado neste artigo. A Fig. 1 apresenta um resumo da pesquisa bibliométrica realizada.

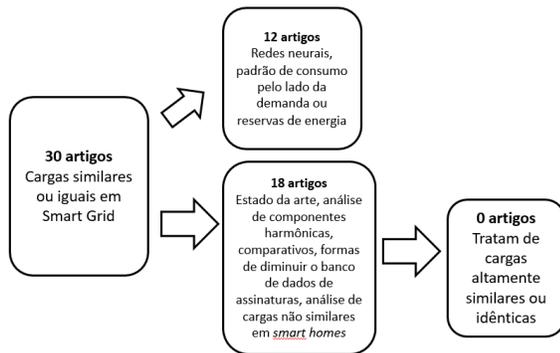


Fig. 1. Pesquisa bibliométrica.

Em [11], [13], [19] e [20] é possível encontrar um ponto comum de que as cargas constituem-se de equipamentos diferentes e que ainda sim são confundidas entre si no classificador quando possuem comportamento e potência parecidos. Os autores ressaltam que as taxas de erro do sistema de identificação crescem à medida que equipamentos com assinaturas semelhantes estão presentes no cenário analisado, reforçando assim a dificuldade em se diferenciar cargas muito parecidas. Em [23] o autor destaca o desenvolvimento de algoritmos robustos que consigam distinguir cargas similares como uma área de pesquisa fundamental para o futuro.

Pela pesquisa feita até então, é observado que a dificuldade em se identificar cargas com alto grau de similaridade é grande e é conhecida, mas ao que se parece, ainda não tem sido devidamente explorada, apresentando-se como um desafio de pesquisa.

### III. METODOLOGIA

#### A. Plataforma Experimental e Dados Utilizados

Quatro lâmpadas fluorescentes compactas do mesmo fabricante e com os mesmos dados técnicos compõem a plataforma experimental. As lâmpadas são tecnicamente idênticas, mas nanometricamente distintas devido, por exemplo, ao processo de fabricação. Por isso, as lâmpadas são chamadas de cargas “similares” e não idênticas.

Para medir a corrente total absorvida pelo circuito, é instalado um sensor de corrente, modelo SCT-013-000, conforme pode ser observado na Fig. 2. Também é utilizado um transformador abaixador para permitir a medição de tensão aplicada às cargas. O transformador tem 300 mA de corrente nominal, três fios no primário (127 ou 220 V), três fios no secundário (+9V ou -9V) e pode ser observado na Fig. 2.

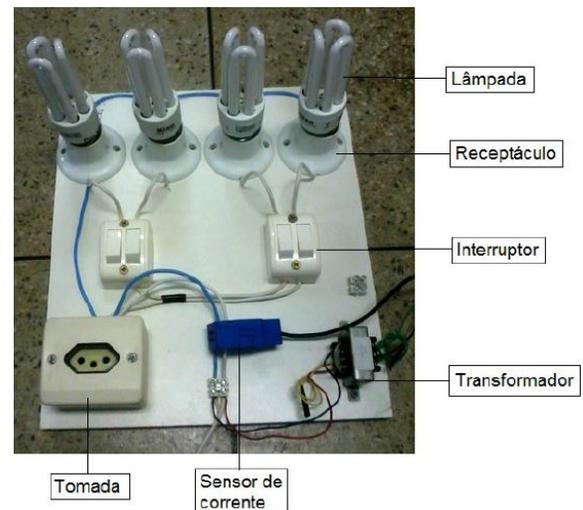


Fig. 2. Componentes da plataforma experimental.

O diagrama esquemático da plataforma é mostrado na Fig. 3, e os dados técnicos das lâmpadas na Tabela I. Um ponto a ser destacado é que a plataforma permite o acionamento individual e combinado das lâmpadas. Sendo assim, é possível ligar as lâmpadas de 16 formas diferentes indo desde todas as lâmpadas desligadas até todas as lâmpadas ligadas. Esse ponto é importante, pois não é analisado apenas o comportamento individual das cargas, mas sim todas as possíveis combinações entre elas.

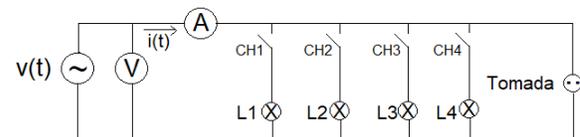


Fig. 3. Esquemático da plataforma experimental.

A plataforma experimental é então energizada e são coletados 150 ciclos de tensão e 150 ciclos de corrente de cada uma das 16 possíveis configurações de lâmpadas. Os 150 ciclos, para cada configuração, são coletados após tempos diferentes de funcionamento do circuito. Dessa forma, considera-se a influência dinâmica do sistema (por exemplo,

TABELA I  
DADOS TÉCNICOS DAS LÂMPADAS

Parâmetro	Valor
Potência ativa	15 W
Tensão nominal	127 V
Corrente elétrica	190 mA
Fator de potência	$\geq 0,5$
Eficiência luminosa	56,3 lm/W
Temperatura máxima do invólucro	85 °C
Faixa de temperatura ambiente	5 a 45 °C
Frequência nominal de alimentação	50/60 Hz
Tempo para atingir 80% do fluxo luminoso	60 s
Fluxo luminoso	844 lm
Vida mediana	6000 h

efeitos térmicos). Especificamente, foram coletados 30 ciclos ao ser ligado, 30 ciclos após 5 minutos de funcionamento, 30 ciclos após 10 minutos, 30 ciclos após 15 minutos e 30 ciclos após 20 minutos de funcionamento. Cabe enfatizar que em todos os casos supracitados o sinal é coletado em seu estado estacionário. Essa coleta é feita com um osciloscópio *Tektronix TDS2002B* e um *notebook* através do software *NI LabVIEW Signal Express Tektronix Edition*.

A frequência de amostragem utilizada é de 25kHz (a máxima frequência permitida pelo equipamento de aquisição utilizado), sendo necessário, para implementar a primeira técnica de caracterização, os dados de corrente e tensão coletados nessa frequência, enquanto que para a segunda técnica de caracterização bastam apenas os dados de corrente. Desse modo, a matriz de dados que se utiliza na segunda técnica de caracterização é composta por dados da corrente elétrica total do circuito, na frequência de amostragem de 25kHz, com 150 ciclos, coletados em momentos distintos, para cada uma das 16 configurações. Totaliza-se assim, 2.400 ciclos de corrente no banco de dados.

### B. Técnicas de Caracterização de Cargas

Duas técnicas de caracterização de cargas elétricas são comparadas. A primeira técnica consiste em representar a carga, ou o conjunto delas, através de 14 características obtidas por manipulações matemáticas dos dados de corrente e tensão coletados para aquela configuração [10]. As características utilizadas nessa primeira técnica são macroscópicas e microscópicas [9]. Dentre as características macroscópicas utilizadas, estão a corrente elétrica eficaz, a potência aparente, a potência ativa, o fator de potência e a potência reativa. As características microscópicas são as componentes fundamental e as 8 primeiras harmônicas ímpares do sinal de corrente decomposto através da transformada discreta de Fourier.

A segunda técnica de caracterização de cargas utilizada está descrita em [29]. Em [29] são utilizados três coeficientes - entropia de Shannon (SE), entropia de Renyi (RE) e banda espectral de energia (SBE) - para definir a assinatura de carga. Essas características são classificadas como microscópicas [9] e são extraídas do sinal de corrente elétrica coletado pelo sensor. Ou seja, nessa segunda técnica não há necessidade das informações de tensão coletadas.

Para extrair do sinal as características que o representam, considerando a segunda técnica, é preciso executar o passo a passo, descrito a seguir, no sinal de corrente elétrica coletado.

- 1) Escolhe-se a quantidade  $n$  de ciclos que é considerada para a extração das características. Neste trabalho, define-se  $n = 1$ . Como a frequência de amostragem é 25 kHz e a frequência da rede é 60 Hz, cada ciclo tem aproximadamente 417 pontos. Sendo assim, obtém-se o vetor corrente elétrica no domínio do tempo com  $I_{ti}^a$  com  $i = 1, 2, \dots, 417$  elementos,  $t$  se referindo ao domínio do tempo e sendo  $a$  a configuração dentre as 16 configurações possíveis, ou seja,  $a = 1, 2, \dots, 16$ .
- 2) Aplica-se a transformada discreta de Fourier no vetor corrente elétrica, obtendo-se o vetor corrente elétrica no domínio da frequência:  $[I_{ti}^a] \xrightarrow{FFT} [I_{fi}^a]$ , onde  $f$  se refere ao domínio da frequência.
- 3) A largura da banda  $f_b$  e a quantidade de bandas  $n_b$  de frequência são definidas. Os elementos do vetor  $[I_{fi}^a]$  serão divididos em grupos de acordo com a escolha desses valores. Cada banda  $B_j$  tem um intervalo de frequência dado pelo menor valor de frequência ( $l_j$ ) e o maior valor de frequência ( $u_j$ ) onde  $j = 1, 2, \dots, n_b$ . Para exemplificar, define-se 1 ciclo ( $n = 1$ ) e 3 bandas ( $n_b = 1$ ) com larguras de 60 Hz ( $f_b = 60$ ) cada. Sendo assim, o vetor de corrente elétrica no domínio do tempo tem 417 elementos. Ao aplicar a transformada de Fourier o novo vetor terá 417 elementos também, mas no domínio da frequência. Cada elemento do vetor  $I_{fi}^a$  terá uma diferença de 60 Hz de frequência. Dessa forma, o primeiro elemento do vetor  $I_{fi}^a$  está na frequência de 0 Hz, o segundo elemento está na frequência de 60Hz, o terceiro elemento está na frequência de 120 Hz e assim sucessivamente. O exemplo define três bandas de 60 Hz cada, portanto a primeira banda é composta pelo primeiro e o segundo elementos do vetor corrente elétrica no domínio da frequência. A segunda banda é composta pelo terceiro elemento do vetor e a terceira banda é composta pelo quarto elemento do vetor. Nesse exemplo, todos os outros 413 elementos do vetor  $I_{fi}^a$  são desconsiderados, permanecendo apenas os 4 primeiros elementos do vetor  $I_{fi}^a$ .
- 4) Para cada banda  $B_j$  são computados os cinco coeficientes: entropia de Shannon (SE), entropia de Renyi (RE), banda espectral de energia (SBE), declinação de SE (dSE) e declinação de RE (dRE), conforme as equações seguintes:

$$SE_j^a = \sum_{l_j}^{u_j} |I_{fi}^a| (\log_2 I_{fi}^a) \quad (1)$$

$$RE_j^a = \frac{1}{(1-r)} \log \left( \sum_{l_j}^{u_j} |I_{fi}^a|^r \right), \text{ com } r = 2 \quad (2)$$

$$SBE_j^a = \frac{\sum_{l_j}^{u_j} |I_{fi}^a|^2}{\sum_{l_1} |I_{fi}^a|^2} \quad (3)$$

$$dSE_j^a = \tan^{-1}(SE_{j+1}^a - SE_j^a) \quad (4)$$

$$dRE_j^a = \tan^{-1}(RE_{j+1}^a - RE_j^a) \quad (5)$$

- 5) A “assinatura de carga” é composta pelos cinco coeficientes de cada banda, organizados em um vetor linha, onde o número de colunas é igual ao número de características da “assinatura”. Portanto, se for definido, por exemplo,  $n_b = 3$ , cada ciclo será representado por um vetor linha, que é sua “assinatura”, com 15 elementos (características).
- 6) Após realizar as etapas 2 a 5 para todos os 2400 ciclos, os vetores são empilhados compondo uma matriz onde o número de colunas é igual ao número de características. Cada coluna é normalizada de acordo com (6), onde  $x_k$  é uma amostra de uma característica,  $x'_k$  é uma amostra normalizada de uma característica,  $\min(x)$  e  $\max(x)$  são o menor e o maior valor da característica na coluna, respectivamente. A normalização é necessária para os classificadores diferenciarem as configurações entre si.

$$x'_k = \frac{x_k - \min(x)}{\max(x) - \min(x)} \quad (6)$$

em que  $k$  identifica o elemento de cada vetor coluna, com  $k = 1, 2, \dots, 2400$ .

### C. Combinação ou Classificação das Assinaturas

Os métodos de classificação são mantidos os mesmos utilizados em [10], de modo a permitir uma comparação entre os resultados obtidos em [10] com os resultados obtidos neste trabalho. Sendo assim são utilizados cinco métodos de classificação: *K-Means*, duas variações do RBC e duas redes neurais *Multi-Layer Perceptron* (MLP).

O *K-Means* foi proposto por Hartigan [30] e é um algoritmo de agrupamento (*clustering*). O RBC foi proposto por Kolodner [31], sendo utilizado neste trabalho o RBC puro para realizar a análise de similaridade por meio da distância euclidiana, e em conjunto com o *K-Means* (*K-Means+RBC*) para realizar o primeiro passo de agrupamento, sendo a análise de similaridade dentro do agrupamento escolhido feita através da distância euclidiana. A MLP é treinada com um algoritmo de aprendizagem *Back Propagation* [32] e função de ativação Sigmóide. Uma rede possui 4 neurônios na camada escondida (MLP4), enquanto que a outra possui 8 neurônios (MLP8).

Por fim, é necessário medir o desempenho dos classificadores. Para isso, é utilizado o método da Validação Cruzada (*Cross-Validation*). Primeiramente é feito um pré-processamento que embaralha os dados e separa o conjunto de teste (10% dos dados) do restante. O restante dos dados (90%) é dividido em  $k$  grupos que são organizados em: conjunto de

treinamento ( $k - 1$  grupos) e conjunto de validação (1 grupo). Em outras palavras, o conjunto de treinamento é o banco de dados do sistema e o conjunto de validação são as leituras dos sensores. Essas leituras precisam ser comparadas com o banco de dados para identificar quais cargas estão em funcionamento. Este processo de organização é feito  $k$  vezes, alternando a escolha do conjunto de validação, entre os  $k$  conjuntos totais, de modo que o restante ( $k - 1$ ) sempre se torna no conjunto de treinamento. Cada vez que esse processo é feito, é computado a taxa de erro do conjunto de treinamento, candidato a classificador. Ao final, são gerados  $k$  classificadores, sendo o melhor dentre eles determinado através do menor erro sob o conjunto de validação. O conjunto de teste, separado no início, é submetido ao melhor classificador para verificar a taxa de erro, chamada de erro de teste, em um conjunto de dados “desconhecido” durante a validação cruzada. Após todas essas etapas, obtém-se três erros: erro de treinamento, erro de validação e erro de teste. Para comparar os classificadores é utilizado o erro de teste, por isso, apenas ele é mostrado na tabela dos resultados. [33] demonstra que a Validação Cruzada com 10 subconjuntos ( $k = 10$ ) é um bom método para análise de desempenho de classificadores inteligentes aplicados a problemas reais. Portanto, é utilizado neste trabalho  $k = 10$ .

## IV. RESULTADOS E DISCUSSÃO

Foram avaliados dois métodos de caracterização de cargas, um com 14 características macroscópicas e microscópicas e outro que se baseia apenas em características microscópicas. Os resultados obtidos com a aplicação desse primeiro método, chamado aqui de M1, podem ser vistos na Tabela II. Para esse método a menor taxa de erro obtida foi de 27,50% para a técnica de classificação *K-Means+RBC*.

Para a segunda técnica, baseada em características microscópicas e chamada aqui de M2, é importante lembrar que cada banda de frequência definida gera 5 características, isto é, SE, RE, SBE, dSE e dRE. Para obter os resultados posteriormente apresentados, os parâmetros  $n_b$  e  $f_b$  foram variados e as variações são descritas a seguir.

- M2 com 5 características: primeiro testou-se apenas uma banda ( $n_b = 1$ ) com largura de 1020 Hz ( $f_b = 1020$ ). Considerando essas definições e  $n = 1$ , significa que apenas os primeiros 18 elementos do vetor  $I_{fi}^a$  foram utilizados para computar os 5 coeficientes mostrados nas equações 1 a 5. Sendo assim, cada ciclo, dentre os 2.400 presentes no banco de dados, é representado por apenas 5 características que compõem a sua respectiva “assinatura de carga”. Considerando essa largura de banda, foi computado desde a fundamental até a 17ª harmônica, tanto as pares quanto as ímpares. Foi escolhida até a 17ª harmônica, pois na literatura é relatado que até a 15ª harmônica há informações relevantes na onda [34] e [10] cita que até a 17ª harmônica foi verificada informações relevantes.
- M2 com 15 características: nesse caso foram consideradas 3 bandas, cada uma com largura de 62 Hz. Analizando essa largura de banda, o primeiro intervalo de frequência vai de 0 a 62 Hz, o segundo intervalo de frequência vai

de 63 a 125 Hz, o terceiro intervalo de frequência vai de 126 a 188 Hz e assim sucessivamente. Entretanto, foram considerados apenas 3 intervalos de frequência, a saber, o primeiro intervalo (primeira banda), pois contém a fundamental, o terceiro intervalo (segunda banda), pois contém a primeira harmônica ímpar e o quinto intervalo (terceira banda), pois contém a segunda harmônica ímpar. Dessa forma, têm-se 3 bandas, ou seja 15 coeficientes (5 de cada banda), que representam um ciclo analisado.

- M2 com 45 características: aqui foram consideradas 9 bandas com largura de 62 Hz cada. De forma análoga ao item anterior, foram considerados apenas os intervalos que contém a fundamental e as 8 primeiras harmônicas ímpares, totalizando assim 9 bandas. Como cada banda gera 5 coeficientes, cada ciclo foi representado por um vetor linha com 45 elementos. Na literatura, observa-se normalmente que as harmônicas ímpares guardam mais informações que as harmônicas pares, motivando tal escolha para efeito de testes.
- M2 com 85 características: para comparar com os resultados obtidos no M2 com 45 características, foi considerado aqui desde a fundamental até a 17ª harmônica, totalizando 17 bandas com largura de 62 Hz cada. Neste caso, todas as componentes, isto é, tanto as ímpares quanto as pares, foram consideradas.

Na Tabela II podem ser observados os resultados obtidos para cada conjunto de características supracitados e para cada classificador utilizado.

TABELA II  
TAXAS DE ERRO (%) DOS SISTEMAS DE IDENTIFICAÇÃO

Método e características	K-means				
	K-means	RBC	+RBC	MLP4	MLP8
M1 - 14	29,58	27,92	<b>27,50</b>	67,50	46,67
M2 - 5	47,89	<b>43,28</b>	45,38	92,85	87,39
M2 - 15	52,94	<b>44,54</b>	46,64	91,59	86,55
M2 - 45	28,57	<b>25,63</b>	29,83	92,86	94,53
M2 - 85	60,50	<b>53,36</b>	57,14	93,28	94,54

Os resultados em negrito foram as melhores taxas de erro encontradas para cada *setup* considerado. Dentre os resultados, o M2-45 com RBC foi o que apresentou o melhor desempenho, 25,63% de erro.

Quanto aos classificadores, observou-se que as redes neurais (MLP4 e MLP8) conduziram aos piores resultados. Acredita-se que isso se deve ao dimensionamento da rede neural que precisa ser otimizada para cada caso, sendo essa uma proposta de trabalho futuro.

Com relação aos caracterizadores, o método que considera apenas características microscópicas atingiu o melhor resultado encontrado, o que confirma a representatividade das características microscópicas e aponta para o aprofundamento no estudo dessas características que, ao que parece, podem representar cargas de modo unívoco. Apesar do M2-45 ter gerado o melhor resultado, não se pode afirmar qual a combinação de quantidade de bandas e largura delas que gera o melhor resultado. Isso porque os valores testados não representam todo o universo de possibilidades existentes. Sendo assim, é necessário a simulação de todos os valores, ou a aplicação

de um otimizador, visando encontrar a melhor combinação de número de bandas e largura delas. Essa implementação é sugerida como trabalho futuro e também mostrará se as componentes ímpares são, de fato, mais representativas.

Quanto aos valores alcançados na literatura através da análise bibliométrica feita na Seção II, foi encontrado apenas um trabalho [10] que trata da mesma situação, isto é, cargas altamente similares acionadas individualmente ou de modo coletivo. Baseando-se no resultado do trabalho encontrado, cuja menor taxa de erro foi de 27,50%, observou-se uma melhoria de 1,87% na nova técnica de caracterização implementada. Quanto ao que se encontra na literatura para cargas com baixo grau de similaridade, os resultados observados ultrapassam os 90% de acerto. Cabe observar que os valores alcançados aqui são do início da pesquisa de um cenário desafiador e que, apesar de não alcançarem os valores presentes na literatura para cenários com cargas diferentes, são próximos a estes e muito superiores aos resultados de um sistema que escolhesse ao acaso a resposta (1/16 = 6,25% de chance de acerto).

Considerando a configuração que gerou a menor taxa de erro, ou seja, M2-45 com o RBC como classificador, foram analisados em quais casos o sistema de identificação se confundia e apresentava uma solução incorreta. Para isso, cada configuração, dentre as 16 possíveis, foi chamada de classe e foi numerada conforme a Tabela III.

TABELA III  
CONFIGURAÇÕES DE FUNCIONAMENTO E SUAS RESPECTIVAS CLASSES

Classe	Configuração de funcionamento
C0	Todas as lâmpadas desligadas
C1	Apenas a lâmpada L1 ligada
C2	Apenas a lâmpada L2 ligada
C3	Apenas as lâmpadas L1 e L2 ligadas
C4	Apenas a lâmpada L3 ligada
C5	Apenas as lâmpadas L1 e L3 ligadas
C6	Apenas as lâmpadas L2 e L3 ligadas
C7	Apenas as lâmpadas L1, L2 e L3 ligadas
C8	Apenas a lâmpada L4 ligada
C9	Apenas as lâmpadas L1 e L4 ligadas
C10	Apenas as lâmpadas L2 e L4 ligadas
C11	Apenas as lâmpadas L1, L2 e L4 ligadas
C12	Apenas as lâmpadas L3 e L4 ligadas
C13	Apenas as lâmpadas L1, L3 e L4 ligadas
C14	Apenas as lâmpadas L2, L3 e L4 ligadas
C15	Todas as lâmpadas ligadas

Considerando essas classes foi gerada a matriz de confusão do sistema de classificação. Essa matriz detalha as porcentagens de correspondência em cada classe e pode ser observada na Tabela IV.

TABELA IV  
MATRIZ DE CONFUSÃO DO SISTEMA DE IDENTIFICAÇÃO - TAXAS PERCENTUAIS

Classe real	Classe Identificada						Erro por classe (%)
	C1	C2	C4	C5	C7	C8	
C1	89,80	3,07	0,79	5,32	1,03		10,20
C2	29,86	49,57		20,57			50,43
C4	6,82	15,95	31,53			45,70	68,47
C5	19,30		35,95	44,75			55,25
C7	37,73				62,27		37,73
C8				5,91		94,09	5,91

Na Tabela IV a linha corresponde à classe correta, enquanto a coluna se refere a qual classe foi identificada pelo sistema. As classes que não aparecem na Tabela IV são aquelas em que 100% dos casos foram identificados corretamente, ou seja, a classe identificada é igual à classe correta. As classes onde todas as lâmpadas estão ligadas ou desligadas apresentaram 0% de taxa de erro, por exemplo, e por isso não aparecem na Tabela IV.

A maior taxa de erro foi de 45,70% e acontece quando há a classe C4 (apenas a lâmpada L3 ligada). Nesse caso, o sistema identifica incorretamente a C8 (apenas a lâmpada L4 ligada) como a verdadeira. Nota-se que o sistema confunde uma lâmpada ligada quando na realidade é outra lâmpada que está ligada, confirmando a alta semelhança entre as cargas. A segunda maior taxa de erro foi de 37,73% e ocorre quando a classe verdadeira é a C7 (apenas as lâmpadas L1, L2 e L3 ligadas) e o sistema identifica a C1 (apenas a lâmpada L1 ligada) como a verdadeira. Esses casos podem ser corrigidos criando-se uma etapa adicional ao programa, onde o coeficiente SE é comparado antes. Analisando esse parâmetro, observou-se que ele varia bastante com a quantidade de lâmpadas em funcionamento, pois o coeficiente SE da primeira banda (onde está a componente fundamental) está diretamente relacionado à potência ativa absorvida [29]. Os valores do coeficiente SE da primeira banda, considerando a banda com largura de 62 Hz, para cada um dos 2.400 ciclos de corrente podem ser observados na Figura 4.

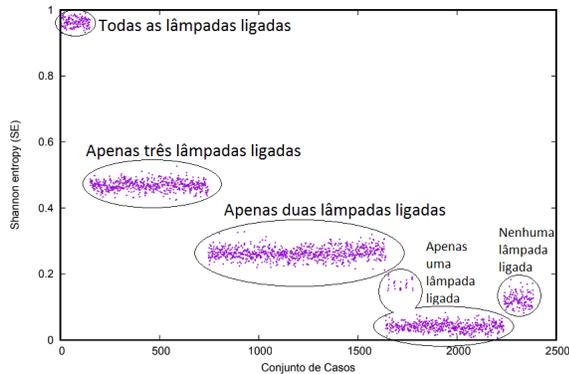


Fig. 4. Valores do coeficiente SE para cada ciclo do banco de dados.

## V. CONCLUSÃO

Com base nos resultados encontrados conclui-se que o método de caracterização baseado em características microscópicas atinge o melhor resultado, cujo valor supera o encontrado no trabalho anterior, avançando o estado da arte em um cenário desafiador e, até então, pouco explorado, que é a identificação de cargas altamente similares. Esses resultados reforçam a utilização das características microscópicas para definir a “assinatura” da carga.

A menor taxa de erro encontrada ocorreu utilizando-se as componentes ímpares do sinal analisado. Esse resultado demonstra que as componentes ímpares podem ser mais representativas e por isso podem definir uma “assinatura” de carga unívoca.

O sistema de identificação, que utiliza características microscópicas, confunde uma lâmpada ligada quando na realidade é outra lâmpada que está ligada, originando a maior taxa de erro do sistema. Isso confirma a dificuldade em se diferenciar cargas altamente similares.

Para diminuir erros onde o sistema de identificação confunde a quantidade de lâmpadas ligadas, pode ser criada uma etapa adicional na classificação onde o valor do coeficiente SE da assinatura é comparado.

Por fim, cabe ressaltar que, ao utilizar características microscópicas, elimina-se a influência do nível de potência da carga. Por isso, o método proposto pode ser aplicado na identificação de cargas similares ou em um conjunto amplo de cargas elétricas distintas. Pode ser observado que C1 é similar a C2, por exemplo, mas não-similar a C1+C2, ou C1+C3, ou ainda a C1+C3+C4, sendo independente do nível de potência da carga. Inclusive a carga C1+C2+C3+C4 tem potência quatro vezes maior que a C3, por exemplo. Sendo assim, o método apresentado pode ser aplicado a cargas similares, ou não, e sua aplicação em cenários com cargas não-similares é um trabalho futuro.

## AGRADECIMENTOS

Os autores agradecem à CAPES/Brasil pela bolsa de mestrado concedida, ao CNPq/Brasil (processo 443384/2014-2) e FAPES-ES/Brasil (Processo .67651259/14) pelo suporte financeiro a esta pesquisa.

## REFERÊNCIAS

- [1] “Electricity statistics. detailed, comprehensive annual data on electricity and heat,” <https://www.iea.org/statistics/electricity/>, accessed: 2019-06-25.
- [2] SANTOS Jr., L. C. N. SOUZA, S. J. SILVA, and W. L. P. SANTOS, *Energias: Reflexões sob a Perspectiva Transdisciplinar*, ser. Gestão de Sistemas de Energia. CRV, 2016, no. 1, p. 13–41.
- [3] K. Ehrhardt-Martinez, K. Donnelly, and J. Laitner, “Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities,” *American Council for an Energy-Efficient Economy Washington, DC*, 2010.
- [4] J. Liang, S. K. K. Ng, G. Kendall, and J. W. M. Cheng, “Load signature study—part i: Basic concept, structure, and methodology,” *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 551–560, 2010.
- [5] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, “A comparative study of lpwan technologies for large-scale iot deployment,” *ICT Express*, vol. 5, no. 1, pp. 1–7, 2019.
- [6] G. W. Hart, “Nonintrusive appliance load monitoring,” in *Proceedings of the IEEE*, ser. 12, vol. 80, 1992, pp. 1870–1891.
- [7] V. P. Borin, C. H. Barriquello, and A. Campos, “Vector projection classification for home appliances recognition: A load signature comparative analysis,” in *Industry Applications (INDUSCON), 2016 12th IEEE International Conference on. IEEE*, 2016, pp. 1–8.
- [8] H. A. D. Azzini, J. F. Fardin, and E. O. T. Salles, “Software para monitoramento não intrusivo de cargas residenciais com fins de auditoria energética,” in *IV Simpósio Brasileiro de Sistemas Elétricos (SBSE)*, 2012.
- [9] M. Zeifman and K. Roth, “Nonintrusive appliance load monitoring: Review and outlook,” *IEEE transactions on Consumer Electronics*, vol. 57, no. 1, 2011.
- [10] A. Paixao, W. C. Celeste, L. O. Rigo Jr, D. Coura, H. Rocha, and S. Rissino, “Classificação inteligente aplicada ao problema de identificação de cargas elétricas ‘idênticas’,” in *Simpósio Brasileiro de Pesquisa Operacional*, ser. Anais do XLVIII SBPO, 2016.
- [11] H. Ahmadi and J. R. Marti, “Load decomposition at smart meters level using eigenloads approach,” *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3425–3436, 2015.

- [12] H. Y. Lam, G. Fung, and W. Lee, "A novel method to construct taxonomy electrical appliances based on load signaturesof," *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 653–660, 2007.
- [13] T. Hassan, F. Javed, and N. Arshad, "An empirical investigation of vi trajectory based load signatures for non-intrusive load monitoring," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 870–878, 2014.
- [14] N. Calamero, Y. Beck, and D. Shmilovitz, "Defining the unique signatures of loads using the currents' physical components theory and z-transform," *IEEE Transactions on Industrial Informatics*, vol. 11, no. 1, pp. 155–165, 2015.
- [15] M. Gonzalez, V. Debussch'ere, and S. Bacha, "A load identification method for residential building applications," in *IEEE International Conference on Industrial Technology (ICIT)*, 2012, pp. 84–88.
- [16] R. Rayudu, C. Tyler, and M. Witherden, "Towards a tool for nonintrusive individual load measurement and identification," in *Innovative Smart Grid Technologies Asia (ISGT), 2011 IEEE PES*, 2011, pp. 1–8.
- [17] E. Gursoy and D. Niebur, "Harmonic load identification using complex independent component analysis," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 285–292, 2009.
- [18] D. He, L. Du, Y. Yang, R. Harley, and T. Habetler, "Front-end electronic circuit topology analysis for model-driven classification and monitoring of appliance loads in smart buildings," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 2286–2293, 2012.
- [19] C. Belley, S. Gaboury, B. Bouchard, and A. Bouzouane, "An eficiente and inexpensive method for activity recognition within a smart home based on load signatures of appliances," *Pervasive and Mobile Computing*, vol. 12, pp. 58–78, 2014.
- [20] N. Sadeghianpourhamami, J. Ruyssinck, D. Deschrijver, T. Dhaene, and C. Develder, "Comprehensive feature selection for appliance classification in nilm," *Energy and Buildings*, vol. 151, no. 15, pp. 98–106, 2017.
- [21] N. F. Esa, M. P. Abdullah, and M. Y. Hassan, "A review disaggregation method in non-intrusive appliance load monitoring," *Renewable and Sustainable Energy Reviews*, vol. 66, pp. 163–173, 2016.
- [22] M. Singh, S. Kumar, S. Semwal, and R. S. Prasad, *Residential Load Signature Analysis for Their Segregation Using Wavelet—SVM*, ser. Kamalakannan C., Suresh L., Dash S., Panigrahi B. (eds) Power Electronics and Renewable Energy Systems. Lecture Notes in Electrical Engineering. Springer, 2015, vol. 326.
- [23] M. Zhuang, M. Shahidehpour, and Z. Li, "An overview of non-intrusive load monitoring: Approaches, business applications, and challenges," in *2018 International Conference on Power System Technology (POWERCON)*, 2018, pp. 4291–4299.
- [24] N. Buneeva and A. Reinhardt, "Ambal: Realistic load signature generation for load disaggregation performance evaluation," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, 2017, pp. 443–448.
- [25] X. Wu, X. Han, and K. X. Liang, "Event-based non-intrusive load identification algorithm for residential loads combined with underdetermined decomposition and characteristic filtering," *IET Generation, Transmission & Distribution*, vol. 13, no. 1, pp. 99–107, 2018.
- [26] W. Dan, H. X. Li, and Y. S. Ce, "Review of non-intrusive load appliance monitoring," in *2018 IEEE 3rd Advanced Information Technology, Electronic and Automation Control Conference (IAEAC)*, 2018, pp. 18–23.
- [27] M. Ma, W. Lin, J. Zhang, P. Wang, Y. Zhou, and X. Liang, "Toward energy-awareness smart building: Discover the fingerprint of your electrical appliances," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 4, pp. 1458–1468, 2018.
- [28] X. Zhang, Z. Luan, and Z. Zhang, "Steady state load decomposition method combining template matching with k-nearest neighbor algorithms," in *2018 International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*, 2018, pp. 250–254.
- [29] A. S. Bouhouras, A. N. Milioudis, and D. P. Labridis, "Development of distinct load signatures for higher efficiency of nilm algorithms," *Electric Power Systems Research*, vol. 117, pp. 163–171, 2014.
- [30] J. A. Hartigan, *Clustering algorithms*. John Wiley & Sons, Inc, 1975.
- [31] J. L. Kolodner, *Case-Based Reasoning*, ser. Morgan Kaufmann Series in Representation & Reasoning. Morgan Kaufmann Pub, september 1993.
- [32] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, 1986.
- [33] R. Kohavi, "Study of cross-validation and bootstrap for accuracy estimation and model selection," in *International Joint Conference on Artificial Intelligence (IJCAI)*, ser. 2, vol. 14, 1995, pp. 1137–1145.
- [34] D. Srinivasan, W. Ng, and A. Liew, "Neural-network-based signature recognition for harmonic source identification," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 398–405, 2006.



**Rayana Kristina Schneider Barcelos** Possui graduação em Engenharia Elétrica e Mestrado em Energia, ambos pela Universidade Federal do Espírito Santo. Atualmente é professora do Instituto Federal de Brasília.



**Wanderley Cardoso Celeste** É doutor em Engenharia Elétrica pela Universidade Federal do Espírito Santo (UFES) desde 2009. É professor do quadro efetivo do Departamento de Computação e Eletrônica da UFES desde 2009 e participa como membro permanente do Programa de Pós-Graduação em Energia da UFES desde 2011.



**Luís Otávio Rigo Júnior** Possui doutorado em Engenharia de Sistemas e Computação pela Universidade Federal do Rio de Janeiro (2011). Atualmente é professor adjunto do Departamento de Computação e Eletrônica (DCEL), UFES. Atua na área de Aprendizado de Máquina, desenvolvendo soluções para problemas em energia e saúde.



**Gisele de Lorena Diniz Chaves** Possui doutorado em Engenharia de Produção pela Universidade Federal de São Carlos (2009). Atualmente é professora adjunta do Departamento de Engenharias e Tecnologia (DETEC), na Universidade Federal do Espírito Santo.