

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO UNIVERSITÁRIO NORTE DO ESPÍRITO SANTO  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENERGIA

RAYANA KRISTINA SCHNEIDER BARCELOS

**IDENTIFICAÇÃO INTELIGENTE DE CARGAS  
ELÉTRICAS SIMILARES EM SMART GRID**

São Mateus  
2018

RAYANA KRISTINA SCHNEIDER BARCELOS

# **IDENTIFICAÇÃO INTELIGENTE DE CARGAS ELÉTRICAS SIMILARES EM SMART GRID**

Dissertação de Mestrado apresentada ao Programa de Pós-Graduação em Energia da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do Grau de Mestre em Energia.

Orientador: Prof. Dr. Wanderley Cardoso Celeste

Coorientador: Prof. Dr. Luis Otávio Rigo Júnior

São Mateus

2018

---

Rayana Kristina Schneider Barcelos  
Identificação Inteligente de Cargas Elétricas Similares em Smart Grid/ Rayana  
Kristina Schneider Barcelos. – São Mateus, ES, 2018-  
54 p. : il. (algumas color.) ; 30 cm.

Orientador: Prof. Dr. Wanderley Cardoso Celeste

Dissertação de Mestrado – Universidade Federal do Espírito Santo, 2018.

1. Identificação de Carga. 2. Assinatura de Carga. 3. Smart Grid. I. Barcelos,  
Rayana Kristina Schneider. II. Universidade Federal do Espírito Santo. IV.  
Identificação Inteligente de Cargas Elétricas Similares em Smart Grid

CDU 02:141:005.7

---

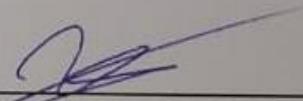
RAYANA KRISTINA SCHNEIDER BARCELOS

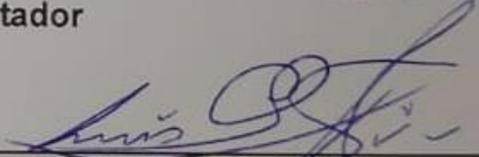
IDENTIFICAÇÃO INTELIGENTE DE CARGAS ELÉTRICAS  
SIMILARES EM SMART GRID

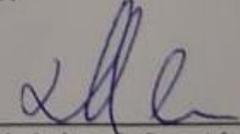
Dissertação apresentada ao Programa de Pós-Graduação em Energia da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Energia.

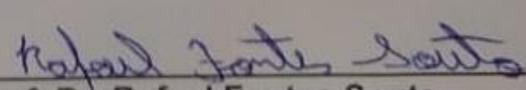
Aprovada em 09 de julho de 2018.

COMISSÃO EXAMINADORA

  
\_\_\_\_\_  
Prof. Dr. Wanderley Cardoso Celeste  
Universidade Federal do Espírito Santo  
Orientador

  
\_\_\_\_\_  
Prof. Dr. Luis Otávio Rigo Júnior  
Universidade Federal do Espírito Santo  
Co-orientador

  
\_\_\_\_\_  
Prof. Dr. Daniel José Custódio Coura  
Universidade Federal do Espírito Santo

  
\_\_\_\_\_  
Prof. Dr. Rafael Fontes Souto  
Instituto Federal de Brasília

*Dedico este trabalho à minha família.*

# Agradecimentos

Agradeço a Deus e à minha família por me apoiarem sempre, bem como aos meus orientadores por todo o suporte dado nessa jornada.

*"It matters not how strait the gate,  
How charged with punishments the scroll,  
I am the master of my fate:  
I am the captain of my soul"*  
*(Invictus)*

# Resumo

Esta dissertação mostra o resultado da aplicação de técnicas de caracterização para definir assinaturas de carga em Smart Grids. O diferencial deste trabalho é que as cargas têm os mesmos dados técnicos e são do mesmo fabricante (cargas com alto grau de similaridade), tornando o processo de identificação mais difícil e descrevendo uma condição desafiadora. O protótipo é uma plataforma com quatro lâmpadas fluorescentes tecnicamente idênticas, permitindo 16 configurações de operação possíveis, isto é, desde nenhuma lâmpada ligada até todas as lâmpadas ligadas. São testadas duas técnicas para definir a assinatura de carga: uma com 14 características simples para representar cada uma das 16 configurações possíveis; e outra forma baseada na entropia de Shannon e Renyi. Cada conjunto de assinatura de carga é submetido a 5 classificadores inteligentes bem difundidos na literatura. Em seguida, os conjuntos de assinaturas, classificados através do Raciocínio Baseado em Casos (RBC), são submetidos a um otimizador objetivando encontrar a maior acurácia possível para o sistema de identificação. A menor taxa de erro obtida neste trabalho foi de 22,69% e representa um bom desempenho do sistema de identificação, tendo em vista a complexidade do problema. Estes resultados iniciais servirão de referência para novas soluções a este novo problema.

**Palavras-chaves:** Identificação de Carga, Assinatura de Carga, Smart Grid.

# Abstract

This work shows the result of applying characterization techniques to define load signatures in Smart Grids. The differential of this work is that the loads have the same technical data and are from the same manufacturer (loads with a high degree of similarity), making the identification process more difficult and describing a challenging condition. The prototype is a platform with four technically identical fluorescent lamps, allowing 16 possible operation configurations, this means, from no one lamp turned on to all the lamps turned on. Two techniques are tested to define the load signature: one with 14 simple features to represent each one of the 16 possible configurations; and another form based on the Shannon and Renyi entropy. Next, the signature sets, classified through Case-Based Reasoning (RBC), are submitted to an optimizer aiming to find the highest possible accuracy for the identification system. The lowest error rate obtained in this work is 22.69% and represents a good performance of the identification system, given the complexity of the problem. These initial results will serve as a reference for new solutions to this new problem.

**Keywords:** Load Identification, Load Signature, Smart Grid

# Lista de ilustrações

Figura 1 – Etapas do processo de identificação de cargas não intrusivo. . . . .	17
Figura 2 – Componentes da plataforma experimental. . . . .	18
Figura 3 – Evolução deste trabalho . . . . .	20
Figura 4 – Pesquisa bibliométrica. . . . .	27
Figura 5 – Diagrama de blocos da metodologia adotada. . . . .	31
Figura 6 – Esquema de chaves que ilustra a metodologia adotada. . . . .	32
Figura 7 – Exemplo da divisão do vetor . . . . .	33
Figura 8 – <i>Cross-Validation</i> . . . . .	37
Figura 9 – Gráfico do número de bandas ímpares e a taxa de erro. . . . .	43
Figura 10 – Convergência do PSO implementado. . . . .	45
Figura 11 – Detalhe do gráfico de convergência do PSO (subida). . . . .	45
Figura 12 – Superfície com todas as taxas de acerto. . . . .	46
Figura 13 – Região do número de bandas onde ocorrem as maiores taxas de acerto. . . . .	47
Figura 14 – Região da largura de bandas onde ocorrem as maiores taxas de acerto. . . . .	47

# Lista de tabelas

Tabela 1 – Comparativo entre as redes convencionais e as <i>smart grids</i> . . . . .	14
Tabela 2 – Comparativo entre diversos sistemas de identificação. . . . .	24
Tabela 3 – Pesquisa bibliométrica. . . . .	28
Tabela 4 – Harmônicas contidas nas bandas de frequência . . . . .	41
Tabela 5 – Taxas de acerto(%) dos sistemas de identificação . . . . .	42
Tabela 6 – Melhoria da taxa de acerto. . . . .	48

# Lista de abreviaturas e siglas

RBC	Raciocínio Baseado em Casos
ILM	Intrusive Load Monitoring
NILM	Non-Intrusive Load Monitoring
PSO	Particle Swarm Optimization
IEEE	Institute of Electrical and Electronics Engineers
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
UFES	Universidade Federal do Espírito Santo

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>14</b>
1.1	Caracterização do problema	15
1.2	Objetivo	18
1.3	Metodologia	18
1.4	Contribuição	20
1.5	Estrutura do trabalho	21
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>22</b>
2.1	Levantamento do Estado da Arte	22
2.2	Pesquisa bibliométrica	26
<b>3</b>	<b>METODOLOGIA</b>	<b>30</b>
3.1	Técnica de caracterização de cargas	32
3.2	Combinação ou classificação das assinaturas	34
3.2.1	K-means	35
3.2.2	RBC	35
3.2.3	Avaliação do desempenho dos classificadores	36
3.3	Otimizador	37
<b>4</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>40</b>
4.1	Varição manual da largura e número de bandas	40
4.2	Otimização	43
4.3	Simulação de todos os valores	46
4.4	Comparação com outros trabalhos	47
<b>5</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>49</b>
5.1	Artigos e trabalhos produzidos	49
5.2	Trabalhos futuros	50
	<b>REFERÊNCIAS</b>	<b>51</b>

# 1 Introdução

A utilização da energia elétrica está associada ao crescimento, ao desenvolvimento tecnológico de um país, ao conforto e à dignidade humana, sendo seu acesso indispensável nos dias atuais e um direito universal. Entretanto, com o crescimento populacional e o aumento do consumo de energia elétrica, acentuado pelo avanço tecnológico, é necessário pensar em formas sustentáveis de produzir e consumir essa energia, seja por motivos financeiros ou pela finitude dos recursos energéticos do planeta (CHAVES; TOSTA, 2016).

O advento da tecnologia da informação e sua aplicação em redes de energia elétrica configura o que é conhecido como rede inteligente ou *smart grid*. Essas redes se utilizam de tecnologia de automação, computação e comunicação na rede elétrica convencional, de modo a obter dados como corrente elétrica, tensão ou demanda da carga, entre outros. Essas informações permitem diversas aplicações, como gerenciamento da energia, auto-recuperação de falhas na rede ou inclusão de equipamentos e consumidores nos processos de planejamento e operação da rede, tornando os sistemas mais confiáveis, seguros e eficientes (FALCÃO, 2010). Na Tabela 1 são listadas algumas diferenças entre a rede convencional e uma *smart grid* (FARHANGI, 2010).

Tabela 1 – Comparativo entre as redes convencionais e as *smart grids*.

<b>Rede convencional</b>	<b><i>Smart Grid</i></b>
Eletromecânica	Digital
Fluxo de comunicação em um sentido	Fluxo de comunicação em ambos os sentidos
Geração centralizada	Geração distribuída
Hierárquica	Rede
Poucos sensores	Sensores em toda parte
Não autônoma	Automonitorada
Restauração manual	Restauração automática
Falhas e Blackouts	Adaptativa e pode operar em ilha
Testes e checagens manuais	Testes e checagens remotos
Controle limitado	Controle universal
Poucas opções de escolha para o consumidor	Muitas opções de escolha para o consumidor

Fonte: Adaptado de (FARHANGI, 2010)(2010)

Algumas diferenças, presentes na Tabela 1, já apontam a evolução das redes convencionais para as redes inteligentes. A popularização do uso de fontes de energias renováveis, como as eólicas e solares, por exemplo, favorece a geração distribuída e, por conseguinte, a implementação de redes inteligentes. O fato das redes inteligentes serem automonitoradas, adaptativas e se restaurarem automaticamente, traz um novo horizonte

para um sistema de proteção mais eficiente, o que também contribui para uma maior utilização de redes inteligentes.

Atualmente, a instalação de medidores inteligentes no sistema de distribuição, conhecidos como *smart meters*, já permite o monitoramento da qualidade da energia fornecida e abre caminho para o controle de demanda, conexão e desconexão de consumidores, ou detecção e isolamento automático de faltas (FALCÃO, 2010).

No que diz respeito à sustentabilidade, o gerenciamento ótimo da energia elétrica, proporcionado pelas *smart grids* em residências, já comprovou uma economia de 10% a 15% da energia utilizada (EHRHARDT-MARTINEZ et al., 2010). Tais vantagens levam a concluir que as *smart grids* podem ser a revolução dos sistemas de fornecimento e consumo de energia elétrica.

Em contrapartida, as diferenças sinalizam também os desafios na consolidação das *smart grids*. O sistema de comunicação, a infraestrutura necessária, como medidores inteligentes, equipamentos telecomandados e sensores, a interoperacionalidade entre as redes, a segurança da informação que circula na rede, tudo isso são aspectos desafiadores, estando atualmente em evolução (FAN et al., 2013).

Há também o problema de identificação das cargas que compõem a rede. Essa identificação é essencial para o gerenciamento das cargas, pois é a partir dela que é possível ligar ou desligar um determinado equipamento, após a análise do seu consumo, por exemplo.

O estudo e desenvolvimento de tecnologia para identificar equipamentos alimentados por uma *smart grid* que seja precisa e robusta pode permitir, dentre outras coisas, o Gerenciamento Pelo Lado da Demanda (*Demand Side Manager - DSM*), o qual é um tema de pesquisa intimamente ligado à questão do uso sustentável dos recursos energéticos disponíveis.

## 1.1 Caracterização do problema

O problema de identificação de cargas pode ser solucionado aplicando-se um sensor em cada carga. Desse modo, é possível saber qual carga está funcionando, ou não, e quanto ela está consumindo.

Esse método é considerado intrusivo, pois necessita de acesso físico às cargas para instalação dos sensores, sendo por isso conhecido universalmente como *Intrusive Load Monitoring* (ILM).

Em uma residência, por exemplo, seria necessário instalar um sensor em cada equipamento a ser monitorado, como chuveiros, tomadas e lâmpadas, o que pode requerer uma infraestrutura que normalmente não está disponível, aumentando o custo de instalação

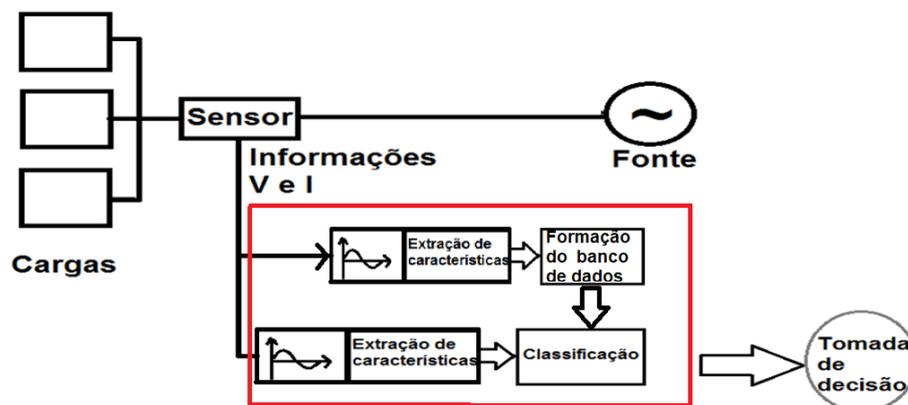
e manutenção da edificação. Por isso, outra metodologia tem sido desenvolvida, onde é instalado apenas um conjunto sensorial na entrada da instalação elétrica a ser monitorada, e, com os dados de medição da demanda elétrica instantânea total, é possível definir quais equipamentos estão ligados e qual seu consumo (instantâneo ou médio). Em uma residência, por exemplo, seria necessário colocar sensores apenas no ramal de entrada, sendo desnecessário o acesso à parte interna da residência. Essa metodologia, proposta no início dos anos 90 por Hart, é conhecida como *Non-Intrusive Load Monitoring*(NILM)([HART, 1992](#)).

O monitoramento não intrusivo requer menos recurso de *hardware* em troca de um *software* mais sofisticado quando comparado com o monitoramento intrusivo ([AZZINI; FARDIN; SALLES, 2012](#)) ([FARINACCIO; ZMEUREANU, 1999](#)). Esse *software* deve ser capaz de desagregar o sinal de entrada coletado de modo a identificar quais cargas estão em funcionamento.

As cargas em questão são equipamentos elétricos e suas características podem ser resistivas, capacitivas, indutivas ou uma combinação dessas. Esses equipamentos também podem operar de formas distintas. Em ([HART, 1992](#)), o autor divide os equipamentos elétricos em três categorias, a saber: ON/OFF, Máquinas de Estado Finitos (MEF) e Continuamente Variáveis. Os equipamentos ON/OFF são aqueles que tem apenas um estado quando estão ligados, alguns exemplos são as lâmpadas, TV's e chuveiros. Já as MEF são equipamentos que possuem um número específico de estados e de transições entre eles, tal qual a máquina de lavar. Por fim, os Continuamente Variáveis possuem um número infinito de estados, sendo os refrigeradores, com inversores de frequência para controle de velocidade, um exemplo desse tipo de equipamento. Independente da categoria, Hart(1992) define o termo assinatura de carga como sendo um parâmetro mensurável que, em meio a carga total, fornece informações sobre a natureza ou estado de operação de uma aplicação individual.

Como um todo, o sistema não intrusivo de identificação de cargas passa por duas etapas distintas. Inicialmente, são extraídas as características do sinal e definida a assinatura de cada equipamento da rede. O mesmo é feito para os casos em que o equipamento funciona juntamente com um ou mais equipamentos do sistema. Essas assinaturas, coletadas previamente, são armazenadas em um banco de dados. A definição das assinaturas, através de características escolhidas, é chamada aqui de caracterização. Em seguida, a assinatura do sinal coletado é comparada com as assinaturas armazenadas previamente, de modo a encontrar a correspondência, entre as assinaturas de entrada e do banco de dados, e identificar o(s) equipamento(s) em funcionamento. Essa segunda etapa é chamada aqui de etapa de classificação. A Figura 1 ilustra as duas etapas que constituem o sistema de identificação.

Figura 1 – Etapas do processo de identificação de cargas não intrusivo.



Fonte: Autora(2018)

A definição da assinatura elétrica de um equipamento é uma peça-chave na identificação das cargas de uma *smart grid*. A escolha dos parâmetros da assinatura influencia diretamente na exatidão do *software* em identificar as cargas que estão em funcionamento. Além disso, a robustez do *software* é testada quando as cargas que compõem a rede têm dados e comportamentos semelhantes. Essa situação pode confundir o *software* e ele pode concluir que um equipamento está ligado, quando na verdade é outro que está.

Na literatura, os autores evitam cenários com cargas semelhantes. De fato, apenas em (PAIXÃO, 2016) que um sistema de identificação foi aplicado em um cenário com cargas altamente similares. Outro ponto a ser destacado é que, na maioria dos trabalhos afins encontrados na literatura, as cargas são acionadas uma de cada vez, facilitando o processo de identificação. Nesse sentido, este trabalho busca um conjunto de características que represente de forma unívoca cargas elétricas tecnicamente idênticas (ou seja, com um alto grau de similaridade) que podem ser acionadas tanto individualmente como coletivamente. Apesar dessas escolhas configurarem um cenário considerado difícil e ainda pouco explorado, é essa a situação encontrada no mundo real, sendo imprescindível submeter os sistemas de identificação a tal cenário, a fim de garantir sua robustez.

Em busca de um sistema preciso e suficientemente robusto, utiliza-se aqui os dados experimentais levantados em (PAIXÃO, 2016), referentes à demanda elétrica de cada uma das 16 configurações possíveis entre quatro lâmpadas fluorescentes que têm as mesmas características técnicas e que são do mesmo fabricante. Vale ressaltar que lâmpadas fluorescentes compactas são equipamentos elétricos do tipo ON/OFF muito comuns em instalações elétricas, sendo este um dos fatores que justificam o seu uso. Além disso, tais equipamentos são de baixo custo, o que facilita a obtenção de dados referentes à demanda elétrica de um número grande de equipamentos altamente similares.

## 1.2 Objetivo

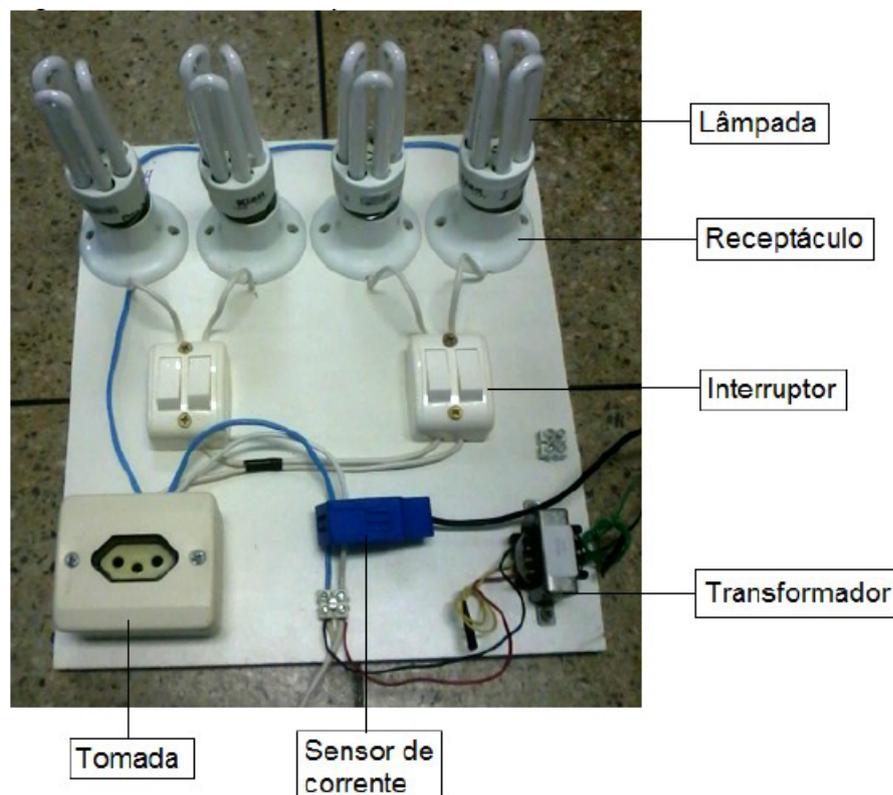
O objetivo deste trabalho é explorar um conjunto de características microscópicas que permita a identificação unívoca de equipamentos elétricos. Para isso, é aplicada uma nova técnica de caracterização, baseada em entropias de Shannon e Renyi, de modo a verificar se os resultados obtidos no sistema de identificação são melhores que os alcançados em (PAIXÃO, 2016).

## 1.3 Metodologia

Este trabalho utiliza os mesmos dados coletados em (PAIXÃO, 2016). Naquele trabalho a plataforma experimental era composta por quatro lâmpadas fluorescentes compactas da mesma marca e com os mesmos dados técnicos. Desse modo, as cargas eram tecnicamente idênticas, porém, devido às diferenças nanométricas causadas, por exemplo, pelo processo de manufatura, elas não eram perfeitamente iguais. Por isso, são aqui chamadas de cargas similares.

A plataforma experimental usada em (PAIXÃO, 2016) para aquisição dos dados experimentais requeridos neste trabalho pode ser observada na Figura 2.

Figura 2 – Componentes da plataforma experimental.



Fonte: Paixão(2016)

É importante notar que a plataforma permite o acionamento individual das lâmpadas. Dessa forma, é possível manter em funcionamento apenas a lâmpada 1, ou apenas as lâmpadas 1 e 2, por exemplo, bem como as lâmpadas 1, 2 e 3 ou todas as lâmpadas. Portanto, tem-se um total de 16 configurações de funcionamento que vão desde todas as lâmpadas desligadas até todas as lâmpadas ligadas. Esse é um ponto a ser ressaltado, pois não é analisado apenas o comportamento individual das cargas, mas sim todas as possíveis ligações entre elas: individuais e combinadas.

Em (PAIXÃO, 2016) é coletado um total de 150 ciclos de tensão elétrica e 150 ciclos de corrente elétrica, oscilando ambas na frequência de 60Hz, para cada uma das 16 possíveis configurações de estado on/off das lâmpadas. Dessa forma, o banco de dados pré-processado de (PAIXÃO, 2016) tem um total de 2.400 ciclos de tensão elétrica e 2.400 ciclos de corrente elétrica. Esses ciclos são obtidos após tempos distintos de funcionamento do circuito para garantir uma maior robustez das amostras coletadas, tendo em vista que os equipamentos elétricos podem exibir pequenas variações ao longo do tempo de funcionamento que podem ser parte de suas características individuais. Sendo assim, de acordo com (PAIXÃO, 2016), dos 150 ciclos de corrente elétrica de cada configuração de funcionamento, 30 são adquiridos após o circuito ser ligado e o sinal estar em estado estacionário, outros 30 ciclos de corrente elétrica são coletados após 5 minutos de funcionamento do circuito, mais 30 ciclos de corrente elétrica são coletados após 10 minutos de funcionamento, outros 30 ciclos de corrente elétrica são coletados após 15 minutos de funcionamento e por fim, mais 30 ciclos de corrente elétrica são coletados após 20 minutos de funcionamento. Os mesmos tempos e quantidade de ciclos são utilizados para compor os 150 ciclos de tensão da configuração de funcionamento. A frequência de amostragem usada em (PAIXÃO, 2016) foi de 25kHz.

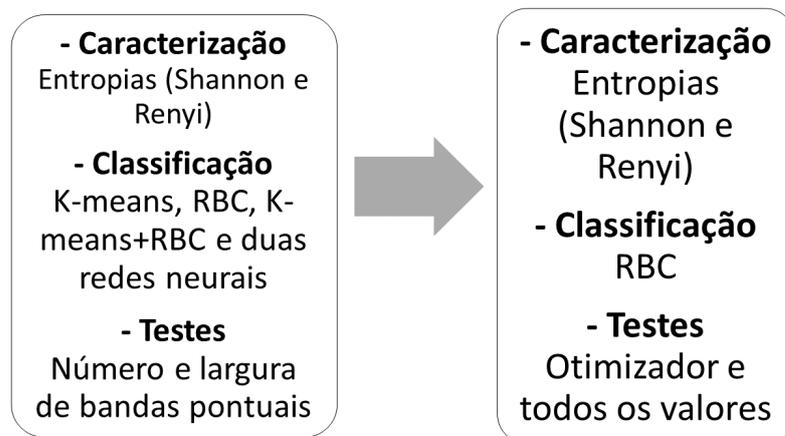
Para caracterizar eletricamente cada carga (lâmpada) da plataforma e cada combinação possível delas, Paixão et al.(2016) representa a carga, ou o conjunto delas, através de 14 características obtidas por manipulações matemáticas dos dados de corrente e tensão coletados para aquela configuração(PAIXÃO, 2016). As características utilizadas são definidas como macroscópicas e microscópicas (ZEIFMAN; ROTH, 2011). Dentre as características macroscópicas utilizadas estão: corrente elétrica eficaz, potência aparente, potência ativa, fator de potência e potência reativa. As características microscópicas são as componentes fundamental mais as 8 primeiras harmônicas ímpares do sinal de corrente decomposto pela transformada discreta de Fourier. Essa técnica de caracterização, presente em (PAIXÃO, 2016), é diferente da implementada neste trabalho.

Para a etapa de caracterização desenvolvida neste trabalho, é implementada uma técnica baseada na entropia de Shannon e Renyi, utilizando apenas o sinal de corrente total do circuito. A etapa de caracterização, baseada nessas entropias, depende da escolha de dois parâmetros que podem variar, sendo eles a largura das bandas e a quantidade delas. Essa técnica é descrita em (BOUHOURAS; MILIOUDIS; LABRIDIS, 2014) e, desse artigo,

apenas a etapa de caracterização é utilizada. A classificação, ou seja, a segunda etapa do sistema de identificação, é inicialmente mantida a mesma de (PAIXÃO, 2016) para fins de comparação. Os métodos de classificação utilizados são o *K-means*, o Raciocínio Baseado em Casos (RBC), uma combinação do RBC com o *K-means* e duas redes neurais, uma com 4 e outra com 8 neurônios na camada escondida. Para esse sistema de identificação são testadas larguras e número de bandas pontuais.

A partir dos primeiros resultados do sistema de identificação, é feita uma análise e é escolhido apenas um classificador, dentre os cinco que (PAIXÃO, 2016) implementou. Nesse segundo momento, o sistema de identificação, com apenas um classificador implementado, passa por um otimizador, baseado no método do enxame de partículas, na busca pela menor taxa de erro do sistema de identificação. Sequencialmente, são simulados todos os valores possíveis da técnica de caracterização implementada, no intuito de certificar-se de que o valor encontrado pelo otimizador é, de fato, a combinação ótima de número de bandas e largura delas. Na Figura 3 podem ser observadas as etapas pelas quais este trabalho passa.

Figura 3 – Evolução deste trabalho



Fonte: Autora(2018)

## 1.4 Contribuição

O caso que se estuda neste trabalho tem como diferenciais, em relação ao que é encontrado na literatura, a identificação de cargas com alto grau de similaridade, além do fato de tais cargas poderem ser ligadas individualmente ou em conjunto, elevando o grau de dificuldade do problema de identificação, porém agregando componentes que aproximam o problema à realidade observada. Portanto, o intuito desse trabalho é avançar no estado da arte ao aplicar ferramentas consolidadas em um cenário ainda pouco explorado e desafiador.

## 1.5 Estrutura do trabalho

No capítulo 2 é mostrado o estado da arte e o procedimento da pesquisa bibliométrica na procura de artigos que tratem de cargas altamente similares. No capítulo 3 é feita uma descrição das características que representam as cargas, dos métodos de classificação aplicados e do otimizador que é implementado. No capítulo 4 é mostrado e discutido o resultado da aplicação das técnicas descritas. No capítulo 5 são comentadas as conclusões e as possibilidades em trabalhos futuros.

## 2 Revisão da literatura

Um dos desafios do monitoramento não-intrusivo é o de desagregar o sinal total coletado pelo sensor, de modo a permitir a identificação de cargas pertencentes a uma base de dados previamente determinada (LIANG et al., 2010). Para isso, é preciso caracterizar cada carga, isto é, buscar em cada equipamento uma assinatura específica e, finalmente, verificar quais assinaturas estão presentes no sinal analisado. Sendo assim, o sistema de identificação passa por duas etapas: primeiro, há a definição e extração das características dos sinais, obtendo assim assinaturas das cargas; depois, na segunda etapa, essas assinaturas são combinadas para identificar as cargas em funcionamento.

### 2.1 Levantamento do Estado da Arte

Uma análise da literatura com foco na primeira etapa mostra que, na busca por essa assinatura, diversas técnicas têm sido testadas: transformada de Fourier, transformada *Wavelet*, cálculo da potência ativa(P), da reativa(Q), da aparente e da admitância. As características escolhidas para representar o sinal analisado podem ser macroscópicas, como valor de tensão(V), corrente(I), fator de potência(FP), dentre outros, ou microscópicas, como as componentes harmônicas (ZEIFMAN; ROTH, 2011). Entretanto, não há ainda um consenso formado sobre quais são as características representativas universais (AZZINI; FARDIN; SALLES, 2012) (ZEIFMAN; ROTH, 2011). As técnicas e a obtenção das assinaturas podem ser feitas tanto no estado transitório como no estado estacionário do sinal analisado, dividindo o estudo em duas vertentes.

Sequencialmente, a segunda etapa compara 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. Nessa segunda etapa, pode-se aplicar técnicas de classificação como redes neurais, *K-means*, Raciocínio Baseado em Casos (RBC), dentre outros métodos, sendo também alvo de estudos para melhoria da acurácia do sistema de identificação como um todo. É importante ressaltar que, para calcular a acurácia do sistema de identificação, é necessário passar por essas duas etapas e que, apesar de serem etapas complementares, este trabalho foca na primeira etapa do processo de identificação: a caracterização através da definição da assinatura da carga.

Após a implementação do sistema de identificação, é necessário aplicá-lo em um cenário composto por diversas cargas elétricas, de modo a medir sua taxa de acerto. Essas taxas variam de acordo com a escolha das características, que representam as cargas, e da técnica que é utilizada na classificação delas. Outro fator que influencia as taxas de acerto é o grau de similaridade entre as cargas. A taxa de acerto tende a ser menor à medida

que a similaridade entre o comportamento e os valores nominais das cargas aumenta. O fato das cargas estarem funcionando individualmente ou em conjunto, no momento da medição, também altera a taxa de acerto do sistema de identificação. Quando mais de uma carga está ligada é mais difícil para o sistema de identificação verificar quais as cargas em funcionamento, pois é necessário desagregar o sinal analisado. Em contrapartida, uma carga mantida ligada por vez, no momento de medição, facilita o processo de identificação, como mostrado em (PAIXÃO, 2016).

Em (HASSAN; JAVED; ARSHAD, 2014) é utilizada a trajetória de tensão e corrente, além da potência ativa, reativa e distorção harmônica total para caracterizarem a carga. Para a etapa de classificação são utilizados quatro algoritmos de aprendizagem, incluindo rede neural artificial e máquina de vetor suporte, do inglês *Support Vector Machine*. Os melhores resultados ultrapassam os 98% e acontecem quando o estímulo adaptativo, *Adaptive Boost*, é utilizado em conjunto com a trajetória de tensão e corrente. O autor cita o uso de cargas com assinaturas altamente similares, entretanto não são detalhadas as cargas utilizadas. A definição de assinaturas similares é feita apenas pelo valor da potência ativa e reativa consumidos pelo cenário analisado. Além disso, o autor limita a análise para cargas com mais de 50W de consumo de potência ativa, de modo a diminuir a quantidade de assinaturas próximas. É interessante observar que o banco de dados utilizado pelo autor é público e pode ser acessado em <http://redd.csail.mit.edu/>. O banco de dados é composto por sinais de corrente e tensão elétricas de diversos equipamentos, comumente encontrado em residências. As residências são monitoradas por vários dias e as cargas monitoradas são separadas em classes como iluminação, refrigerador, lava-louças e microondas, dentre outras. O banco ainda disponibiliza as informações de tensão e corrente elétricas em diferentes taxas de amostragem. Para o problema tratado neste trabalho é necessário que os equipamentos sejam “idênticos”, sendo eles do mesmo fabricante, mesmo tipo e com dados técnicos iguais, por isso não foi utilizado o banco de dados supracitado, uma vez que essas informações não estão disponíveis.

Em (DONG et al., 2012) é escolhida uma janela no tempo, baseada no degrau criado pelo evento de ligar ou desligar um equipamento, e nela são computados valores como variação de potência ativa, reativa ou de corrente elétrica. A sequência com que essas variações ocorrem também é considerada, bem como o tempo de funcionamento do equipamento.

A Tabela 2 mostra a taxa de acerto dos sistemas de identificação citados acima e de outros trabalhos, com suas respectivas referências, indicando também quais os métodos de caracterização e classificação que são utilizados. Na Tabela 2 todos os trabalhos fazem uso de sinais agregados, compostos por mais de um equipamento ligado, além de cenários compostos por cargas não similares.

Tabela 2 – Comparativo entre diversos sistemas de identificação.

Referência	Características	Técnicas de classificação	Amostragem	Taxa de acerto
( <a href="#">HASSAN; JAVED; ARSHAD, 2014</a> )	Trajectoria V-I	Estímulo adaptativo	16,5 kHz	>90%
( <a href="#">DONG et al., 2012</a> )	Potência instantânea	Degraus, sequência, direção e tempo de funcionamento	N/A	>90%
( <a href="#">RUZZELLI et al., 2010</a> )	P, Q, FP, valor médio quadrático e de pico de V e I	Rede neural artificial	1/60 Hz	>84%
( <a href="#">FIGUEIREDO; ALMEIDA; RIBEIRO, 2012</a> )	P, Q e FP	Máquina de vetores de suporte, 5 vizinhos mais próximos	2,5 Hz	>95%
( <a href="#">FARINACCIO; ZMEUREANU, 1999</a> )	P	Reconhecimento de padrões	1/16 Hz	>84%
( <a href="#">CHANG et al., 2012</a> )	Energia no estado transiente	Rede neural artificial	15,4 kHz	>39%
( <a href="#">CHANG et al., 2013</a> )	P e Q	Rede neural artificial e Otimização por enxame de partículas	6 kHz	>92%
( <a href="#">CHANG et al., 2014</a> )	Transformada <i>Wavelet</i> de I	Rede neural artificial	15 kHz	>90%

Fonte: Adaptado de ([AHMADI; MARTI, 2015](#))(2015)

Os valores das taxas de acerto são próximos de 100% na maioria dos casos, como pode ser visto na tabela. Entretanto, existem ainda lacunas no estado da arte, como demonstra outros trabalhos analisados a seguir.

Em ([AHMADI; MARTI, 2015](#)) o autor implementa uma técnica utilizada em reconhecimentos faciais. Enquanto no reconhecimento facial o conjunto de características essenciais, identificadas em um rosto, são chamadas de *eigenface*, o autor chama de *eigenload* o conjunto de características essenciais de cada carga. Cada assinatura de carga é então definida como um vetor de pesos que indica a contribuição de cada *eigenload* para a carga em questão. Apesar do sucesso nas taxas de acerto, que é de 100% para 13 ou mais *eigenloads*, o trabalho divide as cargas em categorias genéricas. Por exemplo, há a categoria equipamentos resistivos, nela pode existir ferro de passar roupas, cafeteira, chuveiro elétrico, ou outros equipamentos resistivos. O programa desenvolvido não distingue um equipamento resistivo do outro, sendo esta uma limitação considerável no gerenciamento de cargas.

Em (BELLEY et al., 2014), o programa implementado é testado em uma casa inteligente com 16 equipamentos. É deixado claro que, para equipamentos com as mesmas características, o sistema de identificação não fará a distinção correta com uma simples modificação no programa. Sendo assim, os 16 equipamentos são conectados de modo que, aqueles com características muito próximas são conectados em fases diferentes do circuito elétrico. O número de fase é uma das características levadas em consideração na representação do dispositivo. Logo, para dispositivos semelhantes, essa característica é a que distingue um equipamento do outro. Os resultados alcançados ultrapassam 95% de acurácia em cenários que simulam atividades diárias e que contemplam mais de um equipamento ligado ao mesmo tempo. Entretanto, os erros ainda sim recaem sobre equipamentos com comportamentos semelhantes, reforçando a dificuldade encontrada em se distinguir equipamentos com comportamentos elétricos parecidos, mas aplicações diferentes.

Em (SADEGHIANPOURHAMAMI et al., 2017) o autor testa o programa implementado em um cenário com cargas distintas acionadas uma a uma. Muitos equipamentos alcançam quase 100% de acerto na detecção, como o microondas e a lâmpada incandescente. O equipamento que apresenta menor taxa de acerto é o aquecedor, pois ele é confundido com o secador de cabelo por ambos serem resistivos e possuírem potências semelhantes. Mais uma vez, a similaridade entre os equipamentos se mostra um desafio para o sistema de identificação.

De modo geral, estes 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 (AHMADI; MARTI, 2015)(BELLEY et al., 2014) (SADEGHIANPOURHAMAMI et al., 2017).

Portanto, após uma revisão exaustiva da literatura científica referente ao tema aqui abordado, foi observado que é comum o uso de cargas que apresentem comportamentos e dados técnicos distintos entre si e que as taxas de acerto desses sistemas se aproxima de 100% na maioria dos casos. Foi detectado ainda que, nas metodologias empregadas por diversos autores, tais cargas são acionadas individualmente, isto é, uma de cada vez, facilitando o processo de identificação, pois não há a necessidade de desagregar o sinal de demanda elétrica. É observada também que a dificuldade em se identificar cargas semelhantes é grande e é conhecida, mas tem sido pouco trabalhada até então, apresentando-se como um desafio para os pesquisadores.

Apesar de se mostrar um cenário desafiador, tais situações precisam ser consideradas a fim de se obter sistemas robustos o suficiente para tratar de situações reais. Sendo assim, a metodologia adotada neste trabalho se baseia no uso de cargas com alto grau de similaridade, acionadas ora individualmente ora em conjunto (como de fato acontece em

condições reais), para tratar do problema de identificação de cargas via NILM.

Uma pesquisa bibliométrica, detalhada na seção 2.2 mostra a inexistência de artigos em inglês tratando do problema de identificação de cargas com alto grau de similaridade. Recentemente, em (PAIXÃO, 2016), foi introduzido tal problema, chamado pelo autor de identificação de cargas “idênticas”. Naquele trabalho, o autor constrói uma plataforma experimental com quatro lâmpadas fluorescentes compactas, as quais podem ser acionadas ora individualmente ora em conjunto (são 16 as configurações de acionamento utilizadas). Dos sinais de demanda, o autor extrai 14 características, algumas macroscópicas e outras microscópicas, representadas em 150 ciclos de cada configuração possível. Por fim, o autor utiliza uma série de ferramentas de classificação para treinar, validar e testar o sistema de identificação proposto.

## 2.2 Pesquisa bibliométrica

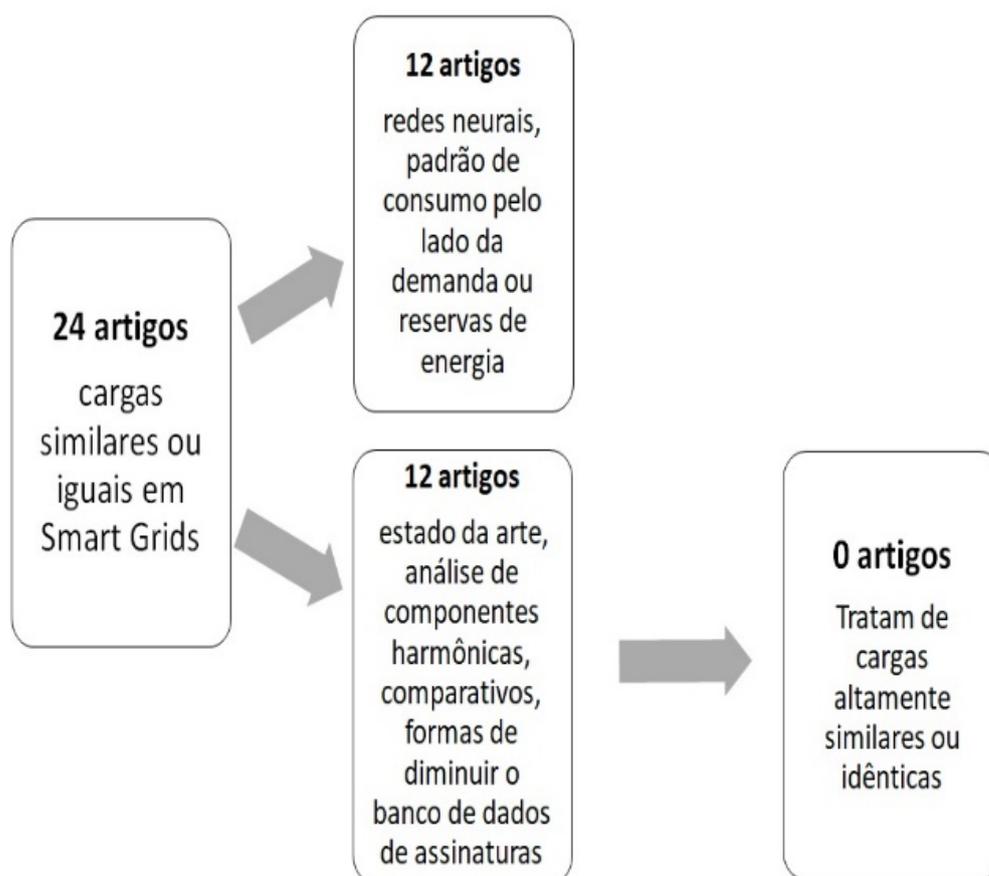
A busca por artigos que tratem de cargas “iguais” ou semelhantes é feita no banco do *Institute of Electrical and Electronics Engineers*(IEEE) e no portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) através de acesso provido pela Universidade Federal do Espírito Santo (UFES) até março de 2018. Ambos são bancos renomados e o portal da CAPES permite a busca em diversos locais como o *Web of Science* e *Elsevier*. Dentre os critérios e tipos de busca, utiliza-se a busca em todo o texto (*full text*) no modo avançado com quatro conjuntos de palavras-chave. É pesquisado um conjunto de palavras-chave por vez e esses conjuntos estão separados por ponto e vírgula conforme segue: “*load signature\**” and “*similar load\**”; “*load signature\**” and “*equal load\**”; “*load identification*” and “*similar load\**” e “*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.

Encontram-se 24 artigos dos quais 12 tratam de assuntos como redes neurais, padrão de consumo pelo lado da demanda ou reservas de energia, não dando enfoque ou tratando de cargas idênticas, de modo que foram retirados após análise. Os doze restantes compreendem trabalhos que mostram o estado da arte, a análise de componentes harmônicas, comparativos entre os diversos métodos de caracterização de cargas existentes ou formas de diminuir o banco de dados de assinaturas. Em nenhum deles são analisadas cargas com um grau de similaridade tão alto como o que é abordado nesse artigo, restando, portanto, 0 artigos que tratem de cargas altamente similares. Em todos os casos, as cargas

constituem-se de equipamentos diferentes e que ainda assim são confundidas entre si no classificador quando possuem comportamento e potência parecidos.

A pesquisa bibliométrica está esquematizada na Figura 4 e detalhada na Tabela 3.

Figura 4 – Pesquisa bibliométrica.



Fonte: Autora(2018)

Tabela 3 – Pesquisa bibliométrica.

Etapa	Atividade	Artigos mantidos
1	Retirada dos artigos repetidos	24 artigos: (KOUZELIS et al., 2015), (CHAKRAVORTY; CHAUDHURI; HUI, 2017), (GUL; MILANOVIC, 2000), (HSU; YANG, 1991), (PARK et al., 1991), (YALLA; KAREEM, 2007), (HASSAN et al., 2013), (SWAMINATHAN; GILBERT, 2012), (VESTRONI; VIDOLI, 2007), (DUAN et al., 2017), (HELMİ; TAYLOR; ANSARI, 2015), (WANG et al., 2016), (AHMADI; MARTI, 2015), (LAM; FUNG; LEE, 2007), (HASSAN; JAVED; ARSHAD, 2014), (CALAMERO; BECK; SHMILOVITZ, 2015), (GONZALEZ; DEBUSSCHÈRE; BACHA, 2012), (RAYUDU; TYLER; WITHERDEN, 2011), (GURSOY; NIEBUR, 2009), (HE et al., 2012), (BELLEY et al., 2014), (SADEGHIANPOURHAMAMI et al., 2017), (ESA; ABDULLAH; HASSAN, 2016) e (SINGH et al., 2015)
2	Retirada de artigos com assuntos diversos como redes neurais ou reservas de energia	12 artigos: (AHMADI; MARTI, 2015), (LAM; FUNG; LEE, 2007), (HASSAN; JAVED; ARSHAD, 2014), (CALAMERO; BECK; SHMILOVITZ, 2015), (GONZALEZ; DEBUSSCHÈRE; BACHA, 2012), (RAYUDU; TYLER; WITHERDEN, 2011), (GURSOY; NIEBUR, 2009), (HE et al., 2012), (BELLEY et al., 2014), (SADEGHIANPOURHAMAMI et al., 2017), (ESA; ABDULLAH; HASSAN, 2016) e (SINGH et al., 2015)
3	Análise dos artigos em relação à tratativa de cargas altamente similares	0

Fonte: Autora(2018)

As características utilizadas para definir a assinatura dos equipamentos podem ser tanto macroscópicas quanto microscópicas, como citado anteriormente. Entretanto, para cargas com alto grau de similaridade, como as que são tratadas neste trabalho, provavelmente as características microscópicas sejam mais representativas. Por exemplo, o DNA ou a impressão digital de um ser humano pode caracterizá-lo de forma unívoca, enquanto que a cor dos olhos, ou do cabelo, não é uma característica forte o suficiente para caracterizá-lo de forma a não ser confundido com outra pessoa. De modo similar, é natural supor que as características microscópicas do sinal analisado serão as responsáveis por representá-lo de modo único.

Essas características microscópicas geralmente advêm de ferramentas de transformação matemática envolvendo sinais, como por exemplo a aplicação da transformada de Fourier e de *Wavelet*. Nessa linha, em (BOUHOURAS; MILIOUDIS; LABRIDIS, 2014)

aplica-se o conceito de entropias de Shannon e Renyi em um cenário com diversas cargas elétricas testadas individualmente. São testados 13 equipamentos diferentes. Sendo 13 equipamentos, testados individualmente, temos então 13 cenários de funcionamento diferentes. Entretanto, alguns equipamentos são testados tanto com 100% da potência de funcionamento, quanto com metade da potência, por exemplo. Isso é feito para equipamentos como o secador de cabelo, em que é possível selecionar o modo quente (100% da potência) ou morno (50% da potência). Dessa maneira, são consideradas então 17 situações diferentes. Destas, nos melhores resultados, apenas uma situação apresentou taxa de acerto inferior a 100%, todas as outras 16 situações são detectadas corretamente em 100% das vezes. Isso chama a atenção e demonstra que as características microscópicas escolhidas podem ser fortemente representativas para os equipamentos elétricos no cenário analisado. Além disso, o sistema de identificação proposto é simples, pois é implementado para trabalhar *online*, adicionando mais uma boa característica ao programa proposto.

Considerando os excelentes resultados, a simplicidade, e que as características microscópicas utilizadas podem ser extremamente representativas, opta-se neste trabalho por utilizar as mesmas características de onda utilizadas em (BOUHOURAS; MILIOLIDIS; LABRIDIS, 2014). Essas características são extraídas e utilizadas para classificar cargas elétricas em um cenário com cargas altamente similares e acionadas individual ou coletivamente, conforme já mencionado.

## 3 Metodologia

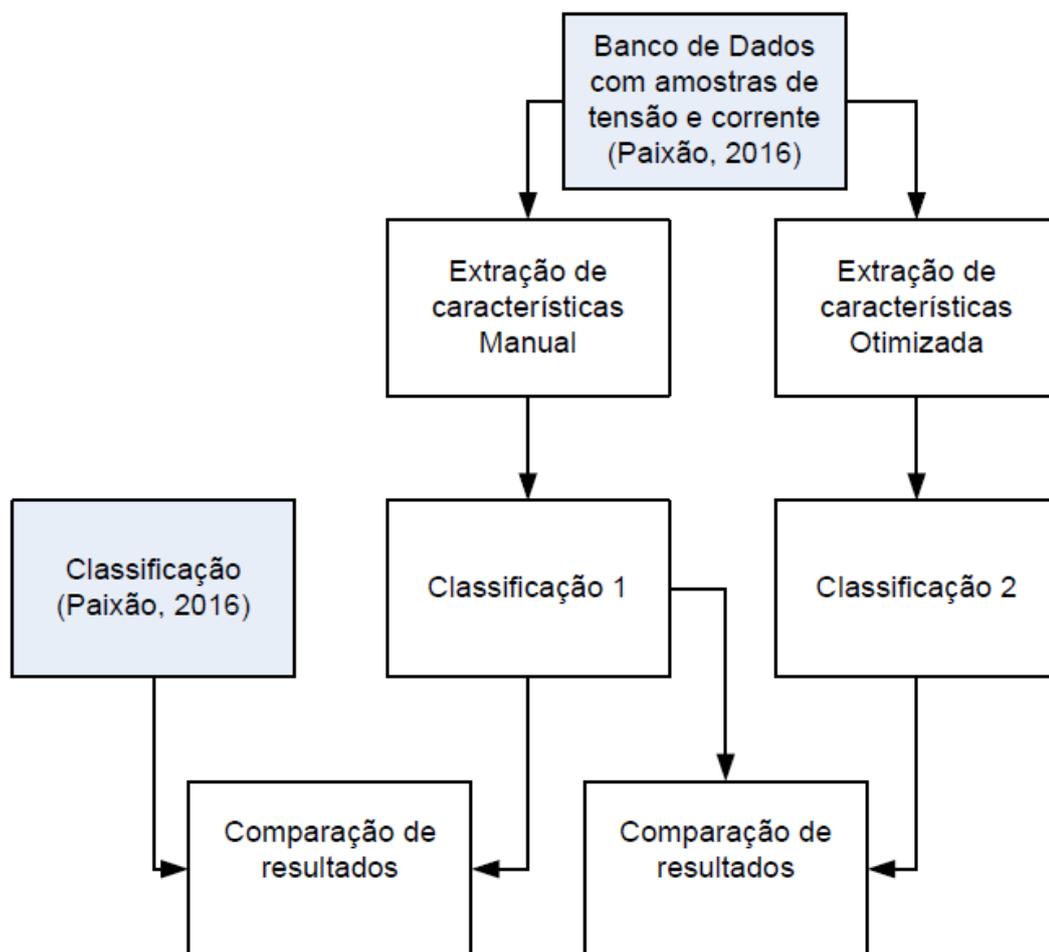
O sistema de identificação que é implementado compõe-se pela parte de caracterização e de classificação. A etapa de caracterização é feita baseada em entropias de Shannon e Renyi. Nessa etapa, é possível variar dois parâmetros, a saber, a largura da banda e a quantidade delas. Cada escolha de número e de largura de bandas gera uma taxa de acerto diferente no sistema de identificação.

A evolução do trabalho que se desenvolve aqui passa por duas etapas distintas. Na primeira, a escolha desses dois parâmetros é feita manualmente e são utilizados os mesmos classificadores de (PAIXÃO, 2016), no intuito de comparar os resultados encontrados com este que é o único trabalho semelhante encontrado na literatura.

Como cada escolha do número de bandas e da largura de cada banda gera uma taxa de acerto diferente, é implementado um otimizador em substituição ao processo manual, a fim de se buscar a combinação que gera a maior taxa de acerto do classificador, sendo esta a segunda etapa deste trabalho. Neste caso, fixa-se o uso de um único classificador dentre os implementados na primeira etapa, a fim de se analisar apenas a influência das características extraídas no processo de identificação. Na segunda etapa também são simulados todas as combinações possíveis entre o número de bandas e a largura delas. O intuito dessa última parte é verificar se o valor encontrado pelo otimizador é um máximo global, e não local.

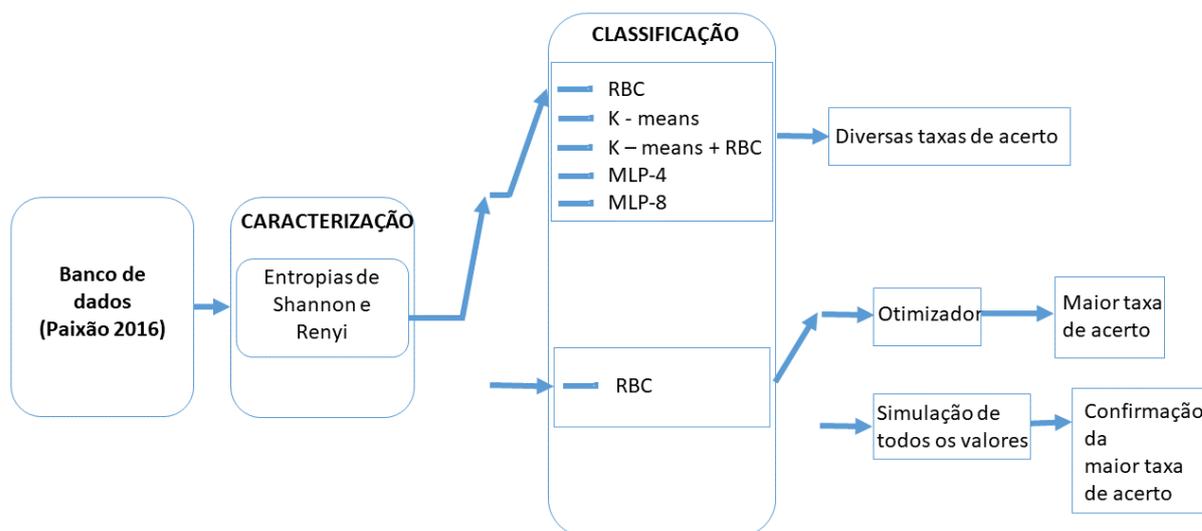
Na Figura 5 é possível observar cada etapa de desenvolvimento do trabalho enquanto que através da Figura 6 é possível visualizar cada etapa do trabalho de modo detalhado.

Figura 5 – Diagrama de blocos da metodologia adotada.



Fonte: Autora(2018)

Figura 6 – Esquema de chaves que ilustra a metodologia adotada.



Fonte: Autora(2018)

### 3.1 Técnica de caracterização de cargas

A técnica de caracterização de cargas usada neste trabalho está descrita no artigo (BOUHOURLAS; MILIOUDIS; LABRIDIS, 2014). Nessa técnica são utilizadas cinco características extraídas dos sinais de demanda dos equipamentos sob identificação, a saber, a entropia de Shannon (SE), a entropia de Renyi (RE), a banda espectral de energia (SBE), a declinação do SE (dSE) e a declinação do RE (dRE). Tais parâmetros, calculados para cada banda, passam a integrar neste trabalho a assinatura do equipamento sob identificação.

A seguir é descrito o passo a passo para a obtenção das assinaturas nessa técnica testada.

1. Define-se a quantidade  $n$  de ciclos da frequência fundamental de alimentação do equipamento que é considerada na extração das características. Para este trabalho,  $n=1$ , isto é, cada característica é extraída de 1 ciclo, assim como foi feito em (PAIXÃO, 2016) e (BOUHOURLAS; MILIOUDIS; LABRIDIS, 2014). Como a frequência de amostragem é de 25 kHz e a frequência da rede é 60 Hz, cada ciclo tem aproximadamente 417 pontos. Desse modo, forma-se o vetor corrente elétrica no domínio do tempo  $I_{ti}^a$  com  $i = 1, 2, \dots, 417$  elementos,  $t$  se referindo ao domínio do tempo e  $a$  sendo 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, isto é  $[I_{ti}^a] \xrightarrow{FFT} [I_{fi}^a]$ , em que  $f$  se refere ao domínio da frequência.

3. São definidas a largura da banda  $f_b$  e a quantidade de bandas  $n_b$  de frequência. Os elementos do vetor  $[I_{fi}^a]$  são 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 60Hz ( $f_b = 60$ ) 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  $i$  desse novo vetor estará em uma frequência específica e a diferença de frequência entre o elemento  $[I_{fi+1}^a]$  e o elemento  $[I_{fi}^a]$  é dada pela Equação 1.

$$\Delta f = \frac{F_s}{n} \tag{1}$$

Onde  $\Delta f$  é a diferença de frequência entre os dois elementos adjacentes,  $F_s$  é a frequência do sinal analisado e  $n$  é a quantidade de ciclos definida na etapa 1. No caso que se estuda neste trabalho, a frequência do sinal é a frequência da rede, ou seja, 60Hz. Portanto, considerando o exemplo onde  $n = 1$  e  $F_s = 60$ , tem-se  $\Delta f = 60$ . Isso significa que 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, então 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. A Figura 7 ilustra essa divisão do vetor, conforme o exemplo colocado.

Figura 7 – Exemplo da divisão do vetor



Fonte:Autora(2018)

4. De cada banda  $B_j$  são obtidas 5 características, ou seja, a entropia de Shannon(SE), a entropia de Renyi(RE), a banda espectral de energia(SBE), a declinação de SE(dSE) e a declinação de RE(dRE). Essas 5 características são calculadas conforme as

Equações (2) a (6), respectivamente.

$$SE_j^a = \sum_{l_j}^{u_j} |I_{fi}^a| \cdot (\log_2 |I_{fi}^a|) \quad (2)$$

$$RE_j^a = \frac{1}{(1-r)} \log\left(\sum_{l_j}^{u_j} |I_{fi}^a|^r\right), r = 2 \quad (3)$$

$$SBE_j^a = \frac{\sum_{l_j}^{u_j} |I_{fi}^a|^2}{\sum_{l_1}^{u_{n_b}} |I_{fi}^a|^2} \quad (4)$$

$$dSE_j^a = \tan^{-1}(SE_{j+1}^a - SE_j^a) \quad (5)$$

$$dRE_j^a = \tan^{-1}(RE_{j+1}^a - RE_j^a) \quad (6)$$

5. A assinatura de carga reúne as 5 características de cada banda formando um vetor, definido como  $\delta_k = [SE_j^a, RE_j^a, SBE_j^a, dSE_j^a, dRE_j^a]$ . Voltando ao exemplo dado na etapa 3, cada uma das três bandas ( $j = 1, 2, 3$ ) gera 5 características, totalizando 15 características. Essas 15 características são a assinatura da carga. É importante observar que o número total de características, ou elementos do vetor assinatura de carga, será sempre o número de bandas vezes 5, uma vez que cada banda gera 5 características. Para este trabalho, o total de ciclos de corrente elétrica no banco de dados é 2.400, onde existem 150 ciclos para cada uma das 16 configurações possíveis de funcionamento da plataforma. Como é utilizado apenas 1 ciclo para representar a carga, cada ciclo do banco de dados gera uma assinatura, logo, para este trabalho,  $k = 1, 2, 3, \dots, 2.400$ .

6. Após realizar as etapas 2 a 5, obtém-se uma matriz  $\delta^{k \times C}$ . Onde  $C$  é o número de características da assinatura de cada carga, dado pela Equação (7).

$$C = 5n_b \quad (7)$$

Além disso, cada coluna da matriz  $\delta$  é normalizada conforme a Equação (8). Isso é necessário para os classificadores diferenciarem as configurações entre si.

$$x'_C = \frac{x_C - \min(x)}{\max(x) - \min(x)} \quad (8)$$

## 3.2 Combinação ou classificação das assinaturas

O intuito deste trabalho é buscar uma melhora na precisão dos sistemas de identificação de equipamentos altamente similares. Para isso, na etapa de caracterização, é testado

um conjunto de características extraíveis dos sinais de demanda elétrica dos equipamentos sob identificação. Para a segunda etapa do sistema de identificação, optou-se por manter 3 dos mesmos métodos de classificação usados em (PAIXÃO, 2016), a saber: *K-means* e duas variações do raciocínio baseado em casos (RBC). Logo, ao variar apenas o conjunto de características que compõem a assinatura dos equipamentos, é possível, através de um processo de comparação entre os resultados obtidos neste trabalho específico e os obtidos em (PAIXÃO, 2016) concluir o quão adequadas são as características aqui escolhidas.

### 3.2.1 K-means

O *K-means* foi proposto por (HARTIGAN; WONG, 1979) e é um algoritmo de agrupamento (*clustering*). Neste trabalho, o algoritmo é utilizado como um aprendizado não supervisionado, em que o agrupamento é realizado por meio da métrica de distância euclidiana, conforme a Equação (9), em que  $E(x,y)$  é a distância euclidiana entre os vetores linha de características  $x$  e  $y$ . Os vetores  $x$  e  $y$  têm  $C$  elementos, e  $r$  identifica o elemento que, neste caso específico, equivale a uma característica.

$$E(x, y) = \sqrt{\sum_{r=1}^C (x_r - y_r)^2} \quad (9)$$

O valor de  $K$  é escolhido automaticamente em função de um fator de entropia médio das classes dentro de cada agrupamento.

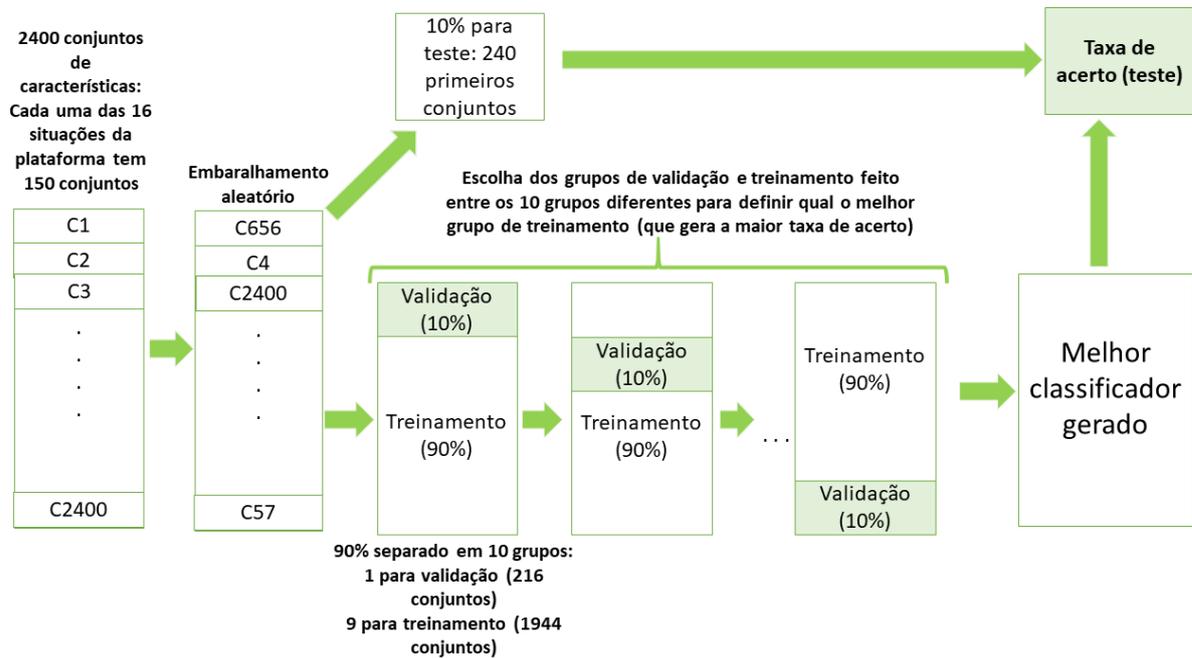
Após o término do algoritmo e formação dos grupos, cada grupo recebe um rótulo de classe que é determinado pela classe majoritária de exemplos dentro do grupo. Ao submeter o classificador a um novo caso, verifica-se o grupo mais próximo deste caso, ou seja, aquele que tem a menor distância euclidiana, e o classifica com o rótulo do grupo.

### 3.2.2 RBC

O RBC foi colocado por (KOLODNER, 2014), e, neste trabalho, são utilizadas duas variações dele. A primeira, denominada RBC, consiste de uma versão simples do algoritmo RBC que efetua a classificação por análise de similaridade através da distância euclidiana entre o caso novo e os exemplos armazenados na base de casos. A segunda variação do algoritmo, denominada *K-means+RBC*, efetua um processo inicial de agrupamento dos exemplos através do algoritmo *K-means* sem a definição de rótulo de classe para o agrupamento. Ao realizar a análise de um caso novo, verifica-se o agrupamento mais próximo deste caso e realiza-se a análise de similaridade dentro do agrupamento escolhido por meio do algoritmo RBC descrito na outra variação.

### 3.2.3 Avaliação do desempenho dos classificadores

O método da Validação Cruzada (*Cross-Validation*) é utilizado para medir o desempenho dos classificadores. Inicialmente é feito um pré-processamento que embaralha os dados e separa o conjunto de teste (10% dos dados, neste trabalho) do restante. Em seguida, o restante dos dados (90%) é dividido em  $k$  grupos. Estes  $k$  grupos são organizados em: conjunto de treinamento ( $k-1$  grupos) e conjunto de validação (1 grupo). 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 transforma no conjunto de treinamento. Esse processo gera  $k$  classificadores, sendo que o melhor dentre eles, determinado através do erro sob o conjunto de validação, é escolhido e guardado, enquanto o erro médio das  $k$  avaliações, sob o conjunto de validação, é usado como resultado do treinamento realizado. Após este processo, o conjunto de teste, separado no início, é submetido ao melhor classificador para verificar a taxa de erro em um conjunto de dados “desconhecido” durante a validação cruzada. Esse processo completo gera três resultados: erro do treinamento, erro da validação e erro do teste. Apenas o erro do teste é utilizado para fins de comparação entre classificadores e, por isso, apenas ele é mostrado na tabela dos resultados. Em (KOHAVI et al., 1995) é demonstrado 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, são utilizados neste trabalho 10 subconjuntos, mesmo valor de  $k$  utilizado em (PAIXÃO, 2016). A Figura 8 ilustra o processo de *cross-validation* para o conjunto de dados e valor de  $k$  utilizados neste trabalho.

Figura 8 – *Cross-Validation*.

Fonte: Autora (2018)

Para definição do grupo de características da carga é utilizado apenas um ciclo, portanto cada ciclo gera uma assinatura de carga. Na Figura 8 os conjuntos de características são representados pela letra C e vão de C1 a C2400. Como o conjunto de dados tem 150 ciclos coletados para cada uma das 16 situações da plataforma experimental, têm-se então 2400 ciclos no total e, por conseguinte, 2.400 assinaturas de carga.

### 3.3 Otimizador

O algoritmo implementado faz a busca pela menor taxa de erro do sistema de identificação da mesma maneira que um bando de pássaros ou um cardume de peixes rastreia sua comida. Esse modelo de otimização é conhecido como *Particle Swarm Optimization*, ou PSO, e foi proposto inicialmente por (KENNEDY; EBERHART, 1995).

A escolha do PSO como método de otimização para este trabalho se deve a uma série de características que se adequam às necessidades. Isto é, o método é de fácil compreensão e implementação computacional, além de se mostrar robusto e de utilizar menos armazenamento de memória e de ter a capacidade de obter soluções globais de qualidade superior a outros métodos como, por exemplo, algoritmos genéticos, com menos tempo computacional (ABBAS et al., 2017). Segundo (ABBAS et al., 2017), o método possui ainda características substanciais de convergência que supera outros métodos de otimização baseados em metaheurísticas.

No PSO, o enxame de partículas é distribuído aleatoriamente num espaço de soluções, sendo que uma das posições das partículas pode ser considerada uma possível solução (KENNEDY; EBERHART, 1995). As partículas têm uma velocidade associada, que indica sua tendência de movimento na próxima interação. Cada partícula do enxame se movimenta no espaço de soluções por meio de duas forças. A primeira é uma atração, de magnitude aleatória, para a melhor posição encontrada por ele mesmo ( $Pbest$ ). A segunda é uma atração, também de magnitude aleatória, para melhor posição encontrada por alguns ou por todos os membros do enxame ( $Gbest$ ). Com essas informações, a posição e a velocidade de cada partícula são atualizadas a cada iteração do algoritmo até o enxame convergir (FIOROTTI et al., 2017). O método do PSO, conforme (FIOROTTI et al., 2017), é descrito a seguir:

**Passo 1:** Iniciar as variáveis do enxame de partículas.

- $y$  representa a posição de uma partícula.
- $U$  é um vetor d-dimensional de números reais aleatórios uniformemente distribuídos entre 0 e 1.
- $y_{min}$  é o menor valor da variável  $y$  que compõe uma possível solução.
- $y_{max}$  é o maior valor da variável  $y$  que compõe uma possível solução.
- $D$  é a dimensão máxima das partículas.
- $y_i$  é a posição de cada componente da partícula.

$$y_i = y_{min} + (y_{max} - y_{min})U, \forall i = 1, \dots, D. \quad (10)$$

- $v_i$  é a velocidade de cada componente da partícula.

$$v_i = (y_{max} - y_{min})U - y_i, \forall i = 1, \dots, D. \quad (11)$$

- $Pbest_i = 0$ , representa a melhor localização da partícula.
- $Gbest = 0$ , representa a melhor localização do enxame.
- $Pbestfun_i = 0$ .
- $Gbestfun_i = 0$ .
- $N$  é o número de partículas (tamanho do enxame).

$$N = 10 + 2\sqrt[2]{D}. \quad (12)$$

- $w$  é o coeficiente de inércia.

$$w = \frac{1}{2\log(2)}. \quad (13)$$

- $c_1$  é o parâmetro cognitivo.

$$c_1 = 0.5 + \log(2). \quad (14)$$

- $c_2$  é o parâmetro social.

$$c_2 = c_1. \quad (15)$$

**Passo 2.** Calcular o valor da função objetivo para todas as partículas.

- O vetor  $Pbest_i$  recebe a posição atual de cada partícula.
- O vetor  $Pbestfun_i$  recebe o valor da função-objetivo da partícula.
- $Gbest$  recebe a posição da melhor partícula do enxame.
- $Gbestfun$  recebe a função-objetivo da melhor partícula do enxame.

**Passo 3.** Atualizar as posições e velocidades das partículas conforme as equações (13) e (14).

$$v_i = wv_i + c_1rand(1, D)(Pbest - y_i) + c_2rand(1, D)(Gbest - y_i) \quad (16)$$

$$y_i = y_i + v_i. \quad (17)$$

**Passo 4.** Calcular a função objetivo  $f(x)$  para todas as partículas.

**Passo 5.** Para cada partícula, comparar o valor da função objetivo atual com o valor de  $Pbest_i$ . Se o valor atual é melhor  $Pbest_i$  recebe a posição atual e  $Pbestfun_i$  recebe o valor da função objetivo atual.

**Passo 6.** Encontrar o melhor valor objetivo entre as partículas atuais e comparar com o  $Gbest$ . Caso haja melhora,  $Gbest$  recebe a posição e  $Gbestfun$  recebe a função objetivo da melhor partícula.

**Passo 7.** Refazer o processo a partir do passo 3 até que uma condição de parada seja encontrada. Cada vez que o processo é executado do passo 3 ao 6 significa que uma iteração ( $P$ ) foi realizada. O critério de parada pode ser definido como um valor específico a ser encontrado pela função objetivo (maior taxa de acerto), ou como um número máximo de iterações a ser realizada pelo programa. O critério de parada adotado neste trabalho é  $P$  atingir o número de iterações máximo definido, uma vez que a maior taxa de acerto (objetivo) é desconhecida.

O uso de cada ferramenta implementada pode ser identificado na Figura 6, onde as chaves de seleção permitem visualizar cada caminho percorrido no trabalho desenvolvido aqui.

## 4 Resultados e discussão

Foram comparados dois métodos de caracterização de cargas, um com 14 características macroscópicas e microscópicas e outro que se baseia apenas em características microscópicas da onda. Para o primeiro método, chamado aqui de M1, o melhor resultado obtido foi uma taxa de acerto de 72,50% para a técnica de classificação RBC+K-*means* e foi desenvolvido no trabalho de (PAIXÃO et al., 2016).

Para a segunda técnica, baseada em características microscópicas e chamada aqui de M2, o trabalho se dividiu em duas etapas. Inicialmente os parâmetros de número e largura das bandas foram variados manualmente, onde pode ser analisada a influência dos classificadores, bem como das componentes pares e ímpares. A segunda parte consistiu na escolha de um classificador e da aplicação do otimizador buscando a maior taxa de acerto para esse método de caracterização. Ainda na segunda etapa, todas as combinações de largura e número de bandas foram testadas, de modo a confirmar a taxa de acerto do otimizador. Nas seções seguintes essas etapas são detalhadas.

### 4.1 Variação manual da largura e número de bandas

É importante lembrar que cada banda de frequência definida gera 5 características: SE, RE, SBE, dSE e dRE. Para obter os primeiros resultados, os parâmetros  $n_b$  e  $f_b$  foram variados manualmente e os valores são descritos a seguir.

1. M2 com 5 características: primeiro testou-se apenas uma banda ( $n_b = 1$ ) com largura de 1020 Hz ( $f_b=1020$ ), sendo assim, foi selecionado desde a fundamental até a 17ª harmônica, tanto as componentes pares quanto as ímpares. Foi escolhida até a 17ª harmônica, pois há na literatura dois trabalhos que relatam que até a 15ª ou 17ª harmônica há informações relevantes na onda (SRINIVASAN; NG; LIEW, 2006) e (PAIXÃO, 2016). Para essa configuração, obteve-se um conjunto de apenas 5 características para representar a carga ou suas combinações. Ou seja, apenas 5 características definem quando a lâmpada 1 está ligada, outras 5 características descrevem quando as lâmpadas 1 e 3 estão ligadas, por exemplo, e assim é feito para os 150 ciclos de todas as 16 possíveis situações.

2. M2 com 15 características: nesse caso foram consideradas 3 bandas, cada uma com largura de 62 Hz. Foi utilizado 62Hz, pois a rede pode não oscilar perfeitamente em 60Hz. Portanto, foi colocado um valor superior para garantir que a componente fundamental do sinal esteja na primeira banda. Aqui foram consideradas apenas a fundamental (contida na primeira banda, de 0 a 62 Hz) e os dois intervalos de frequência que continham as duas primeiras harmônicas ímpares, ou seja, de 125 a 186Hz e de 249 a 310Hz, respectivamente.

A Tabela 4 detalha as bandas, a frequência mínima da banda ( $l_j$ ), a frequência máxima da banda ( $u_j$ ), a largura das bandas e a harmônica contida no intervalo de frequência.

Tabela 4 – Harmônicas contidas nas bandas de frequência

Banda	$l_j$ (Hz)	$u_j$ (Hz)	Largura da banda (Hz)	Componente harmônica contida na banda
1	0	62	62	Fundamental
2	63	124	62	2 <sup>a</sup>
3	125	186	62	3 <sup>a</sup>
4	187	248	62	4 <sup>a</sup>
5	249	310	62	5 <sup>a</sup>

Fonte: Autora(2018)

Para o M2 com 15 características, apenas as bandas 1, 3 e 5, bem como as 5 características de cada uma dessas bandas, foram contabilizadas. Desse modo, a assinatura de carga tem 15 características. Essas bandas foram escolhidas para aproximar o número de características deste trabalho, ou seja, 15, com o número de características utilizado no trabalho de (PAIXÃO et al., 2016), ou seja, 14. É importante notar que, as declinações de SE e RE dependem, respectivamente, do valor de SE e RE da banda subsequente. Isso pode ser observado na Equação (5). Quando apenas as bandas ímpares são consideradas, a banda consecutiva à banda 1, por exemplo, é a banda 3, pois a banda 2 é desconsiderada. Portanto, para calcular o dSE da primeira banda ( $dSE_1$ ), é subtraído o SE da primeira banda ( $SE_1$ ) do SE da terceira banda ( $SE_3$ ), ou seja,  $dSE_1 = SE_3 - SE_1$ . Analogamente, o mesmo raciocínio se aplica para obter o dRE e as declinações das outras bandas ímpares, quando as componentes pares são desconsideradas.

3. M2 com 45 características: aqui foram consideradas 9 bandas com largura de 62 Hz, contabilizando apenas as componentes ímpares desde a fundamental até a 17<sup>a</sup> harmônica. Esse valor foi escolhido para verificar quão forte é a representatividade das componentes harmônicas ímpares. A divisão das bandas segue o mesmo raciocínio de M2 com 15 características, sendo que para o M2 com 45 características serão consideradas apenas as bandas 1, 3, 5, 7, 9, 11, 13, 15 e 17.

4. M2 com 85 características: para comparar com os resultados obtidos no M2 com 45 características foi considerado aqui da fundamental até a 17<sup>a</sup> harmônica totalizando 17 bandas com largura de 62 Hz cada. Neste caso todas as componentes, ímpares e pares foram consideradas.

Na Tabela 5 podem ser observadas as taxas de acerto alcançadas pelos dois os métodos, inclusive o método de (PAIXÃO et al., 2016), chamado aqui de M1, bem como para cada classificador.

Tabela 5 – Taxas de acerto(%) dos sistemas de identificação

Caracterização Método e Características	Classificação		
	K-means	RBC	RBC +k-means
M1-14	70,42	72,08	<b>72,50</b>
M2-5	52,11	<b>56,72</b>	54,62
M2-15	47,06	<b>55,46</b>	53,36
M2-45	71,43	<b>74,37</b>	70,17
M2-85	39,5	<b>46,64</b>	42,86

Fonte: Autora(2018)

Os resultados em negrito destacados na Tabela 5 são as melhores taxas de acerto encontradas para cada *setup* experimentado. Dentre os resultados, o método 2, que considerou a fundamental mais as 8 primeiras harmônicas ímpares, foi o que apresentou o melhor resultado com 74,37% de acerto.

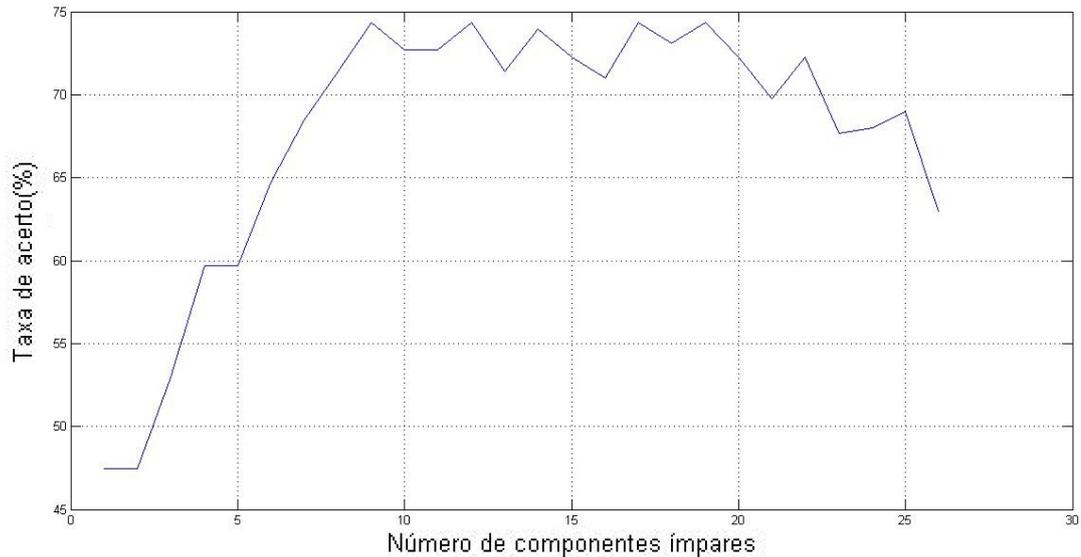
Considerando o método implementado neste trabalho, chamado de M2, o melhor resultado foi obtido com o M2-45, seguido do M2-5. Por exemplo, se a coluna do classificador *K-means* for observada, o *ranking* do melhor para o pior conjunto de características será M2-45, M2-5, M2-15 e M2-85. Se essa mesma linha de raciocínio for aplicada na coluna dos outros dois classificadores, o *ranking* será o mesmo. Ao que parece, se for fixado um método de classificação, as conclusões sobre qual é o melhor ou o pior conjunto de características serão as mesmas.

Com relação aos caracterizadores, o método que considera apenas características microscópicas (M2) atingiu um resultado melhor que o encontrado em M1, 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(o método da “força bruta”) 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.

Primeiramente, foram calculadas todas as opções que consideram apenas as harmônicas ímpares, ou seja, a largura da banda foi fixada em 62Hz e o número de bandas ímpares variou de 1 a 26, sendo o número de bandas um número inteiro. Para exemplificar, quando o número de bandas for 4, significa que é considerada a fundamental (60Hz) mais as 3 primeiras harmônicas ímpares ( $3^a$ ,  $5^a$  e  $7^a$ ). O melhor resultado encontrado foi de 74,37% para 9 bandas e esse valor se repetiu para 12, 17 e 19 bandas. Fora dessa faixa de valores, a taxa de acerto diminui. A curva que relaciona o número de bandas ímpares e a

taxa de acerto do sistema pode ser vista na Figura 9. Como pode ser observado, existe uma faixa de valores, entre 9 e 19 bandas, em que os resultados apresentados são melhores e, fora dessa faixa, eles pioram.

Figura 9 – Gráfico do número de bandas ímpares e a taxa de erro.



Fonte: Autora(2018)

Para considerar as bandas ímpares e pares, devido o universo ser maior, foi implementado um otimizador e os resultados alcançados são mostrados na próxima seção.

## 4.2 Otimização

Na etapa detalhada na seção anterior, há diferenciação entre componentes pares e ímpares, devido a fixação da largura da banda próxima a 60Hz. Para essa segunda etapa, a largura da banda é variada no otimizador, de modo que perde o sentido falar de componentes pares e ímpares. Por exemplo, se a largura de banda for 200Hz, a primeira banda vai conter a fundamental, a segunda e a terceira harmônicas juntas. Portanto, para essa segunda etapa, foram contabilizadas tanto as harmônicas pares quanto as ímpares.

Quanto ao classificador, foi observado que sua escolha parece não afetar o *ranking* de caracterizadores. Para fins de comparação com o único trabalho encontrado na literatura, conforme o capítulo 2 mostrou, foi escolhido um dentre os cinco classificadores utilizados por (PAIXÃO, 2016). Como o classificador RCB mostrou o melhor resultado encontrado para a técnica implementada aqui, ele foi o utilizado nos experimentos com o sistema de otimização automática do conjunto de características.

Nesta etapa do trabalho, dois parâmetros podem variar, isto é, a largura da banda e a quantidade delas, sendo essas as duas dimensões que constituem a posição  $y$  da partícula.

O passo que é utilizado neste trabalho, tanto para o número de bandas quanto para a largura de bandas, é 1. Sendo assim, ambos sempre são números inteiros. Para o PSO, foram definidos os seguintes valores:

Largura mínima da banda: 55Hz

Largura máxima da banda: 1200Hz

Número mínimo de bandas: 1

Número máximo de bandas: 17

Número máximo de avaliações do PSO (critério de parada): 400

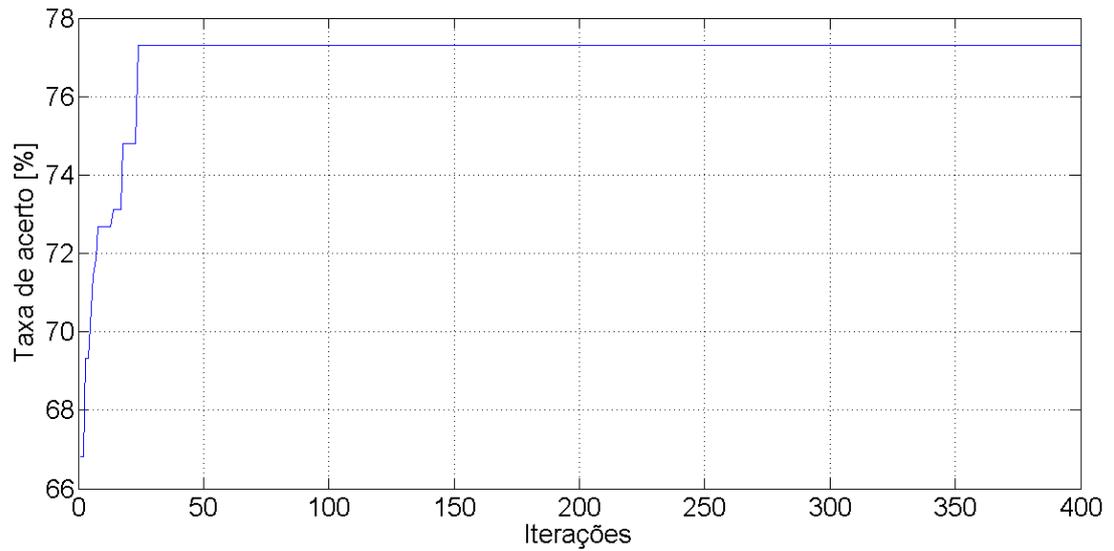
Dimensão do problema: 2

Tamanho do enxame: 14

As larguras consideram desde a fundamental até a vigésima harmônica. O número de bandas permite ir, também, desde a fundamental até a 17<sup>a</sup> harmônica.

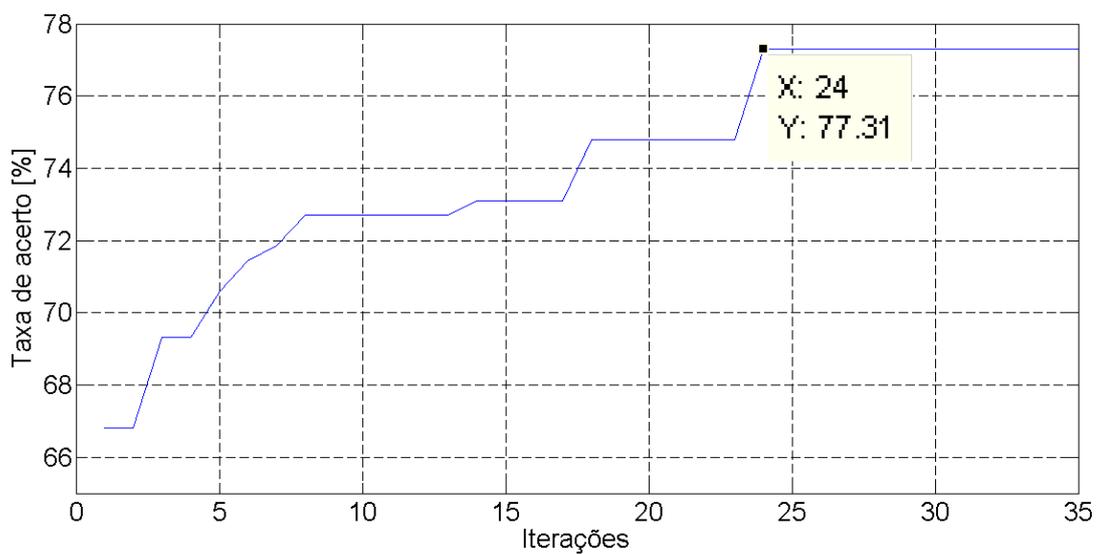
Para validar o PSO é necessário verificar a sua convergência. Essa convergência é observada através de um gráfico, onde o eixo X contém o número da iteração e o eixo Y contém a taxa de acerto obtida naquela iteração. O número de iterações neste trabalho é representado por  $P$  e foi definido como 400. Nas primeiras iterações, as taxas de acerto tendem a aumentar de uma iteração para a próxima iteração, indo em direção à maior taxa de acerto. Entretanto, após algumas iterações iniciais, o PSO alcança a maior taxa de acerto e as novas iterações não mudam mais a taxa de acerto, estabilizando o valor encontrado. Quanto menor o número de iterações necessárias para atingir a maior taxa de acerto, mais rápida é a convergência do PSO implementado. E em quanto mais iterações o PSO encontrar a mesma taxa de acerto máxima, maior o indicativo de que aquele valor é, de fato, a maior taxa de acerto possível. Através da Figura 10, é possível verificar a convergência do otimizador implementado. O detalhe da subida presente no gráfico da Figura 10, pode ser visto na Figura 11.

Figura 10 – Convergência do PSO implementado.



Fonte: Autora(2018)

Figura 11 – Detalhe do gráfico de convergência do PSO (subida).



Fonte: Autora(2018)

Com os valores definidos anteriormente e utilizando o classificador RBC, o PSO encontrou uma taxa de acerto de 77,31%, que é maior que a taxa de acerto encontrada na etapa anterior com M2. Além disso, o PSO convergiu para a maior taxa de acerto já na 24ª iteração, como detalhado na Figura 11.

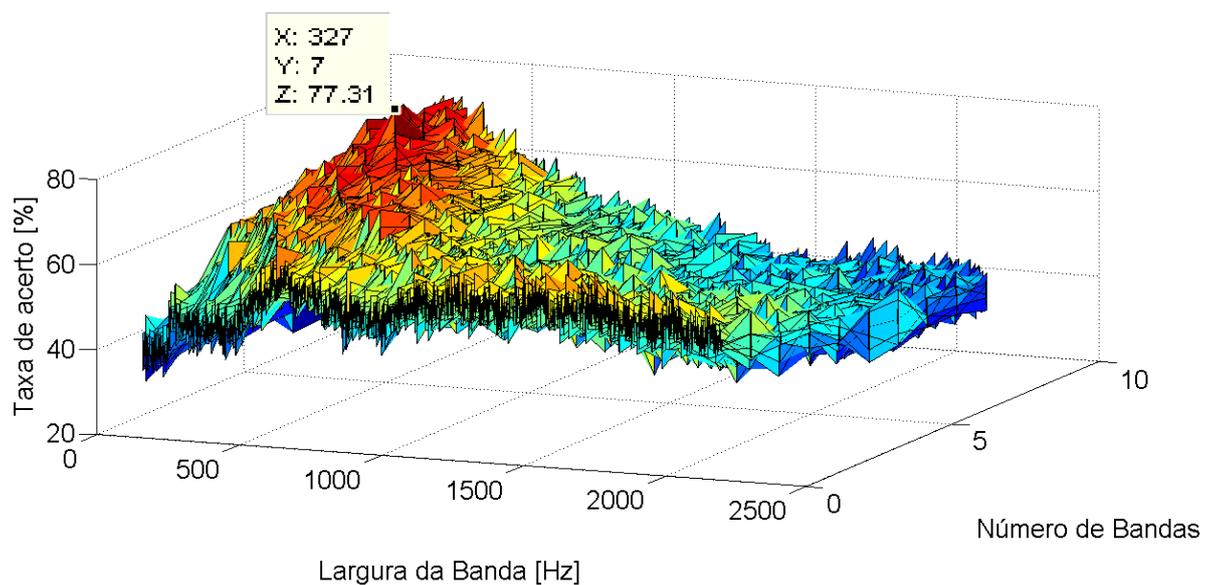
O valor ótimo de 77,31% foi encontrado para 7 bandas com 327 Hz de largura cada e é importante observar o comportamento do número de bandas e sua largura. A maior taxa de acerto fica em 7 bandas com 327 Hz de largura cada, o que significa que são 2.289

Hz de frequência analisados. Essa faixa de frequência se estende desde a fundamental até a 38ª harmônica, sinalizando que para a análise de cargas similares, é necessário observar componentes para além da 17ª harmônica. Além do mais, ao que parece, as componentes pares e ímpares juntas fornecem dados mais importantes que apenas as componentes ímpares, uma vez que o melhor resultado alcançado foi obtido com componentes pares e ímpares.

### 4.3 Simulação de todos os valores

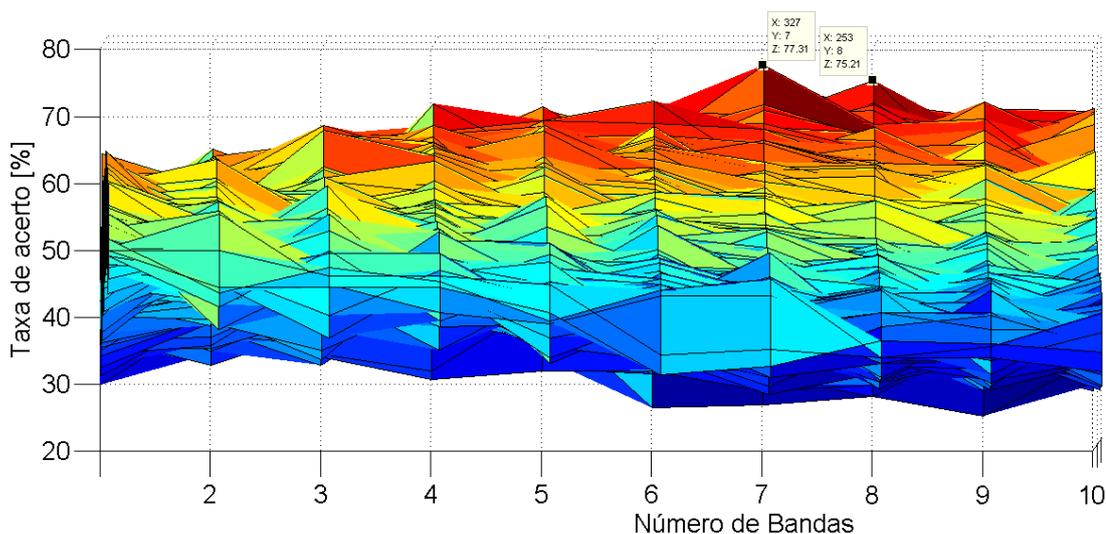
Mesmo com a utilização do otimizador, ainda havia dúvida quanto ao valor encontrado ser o maior possível. Isso se deve ao fato de que o otimizador pode cair em um máximo local e não em um máximo global. Para eliminar tal dúvida, foram simulados todos os valores com largura de banda indo de 60 a 2100 Hz, de 1 em 1 Hz, e a quantidade de bandas igual a 1, 2, 3, ..., 10. Dessa forma, foi obtida uma superfície que pode ser vista na Figura 12, onde o maior valor encontrado foi de 77,31%, com 7 bandas de 327 Hz de largura.

Figura 12 – Superfície com todas as taxas de acerto.



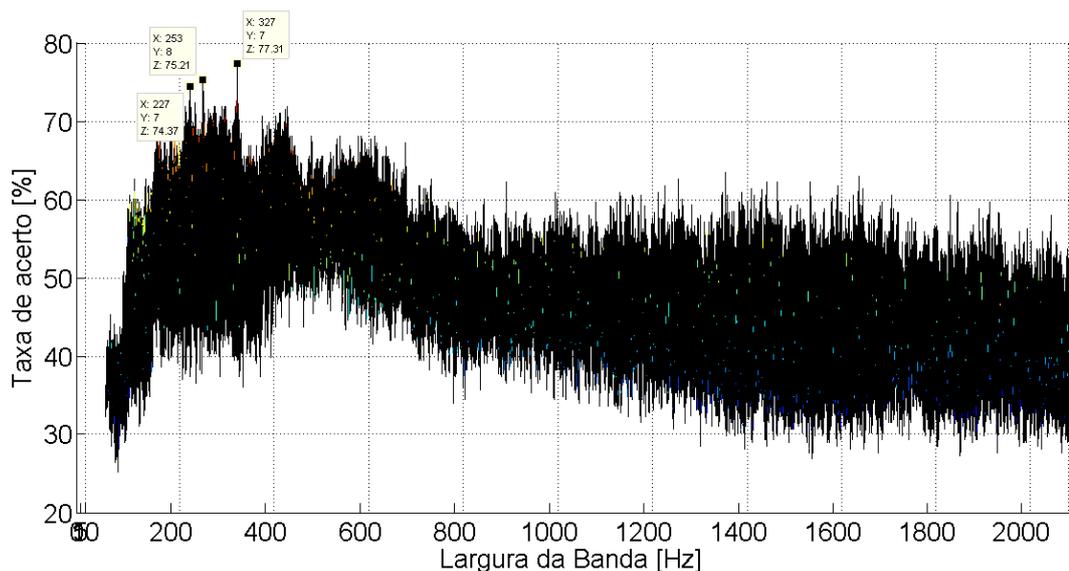
O valor encontrado confere com o valor colocado anteriormente. Um ponto de interesse é a região onde as maiores taxas de acerto aparecem. Para visualizar melhor essa região, foram geradas as Figuras 13 e 14, de modo a evidenciar seus limites.

Figura 13 – Região do número de bandas onde ocorrem as maiores taxas de acerto.



Fonte: Autora(2018)

Figura 14 – Região da largura de bandas onde ocorrem as maiores taxas de acerto.



Fonte: Autora(2018)

Como pode ser observado, entre 6 e 8 bandas e entre 220 e 330 Hz ocorrem as maiores taxas de acerto.

#### 4.4 Comparação com outros trabalhos

Quanto aos valores alcançados na literatura através da análise bibliométrica feita no capítulo 2, foi encontrado apenas um trabalho (PAIXÃO et al., 2016) que trata da mesma

situação, isto é, cargas altamente similares em funcionamento individualmente ou de modo coletivo. Baseando-se no resultado do trabalho encontrado, cuja maior taxa de acerto foi de 72,50%, observou-se um aumento para 77,31% na nova técnica de caracterização implementada. Essas informações estão resumidas na Tabela 6.

Tabela 6 – Melhoria da taxa de acerto.

Tipo de características	Taxa de acerto obtida(%)
Macroscópicas e microscópicas (PAIXÃO et al., 2016)	72,50%
Microscópicas (variação manual)	74,37%
Microscópicas (otimizador)	77,31%
Microscópicas (todas as combinações)	77,31%

Fonte: Autora(2018)

Quanto ao que se encontra na literatura para cargas com baixo grau de similaridade, os resultados observados ultrapassam os 90% de acerto conforme relatado no capítulo 2. Cabe observar, entretanto, 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).

## 5 Considerações Finais

Com base nos resultados encontrados, conclui-se que os métodos aplicados produzem resultados satisfatórios para o cenário analisado e considerando ainda os resultados apresentados até então na literatura para tal cenário. O método 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. Além disso, a fixação de um método de classificação não causa impacto no conjunto de características escolhido, o que leva a concluir que a fixação de um classificador que apresenta um bom desempenho é uma boa estratégia para a busca e refinamento de novos conjuntos de características que sirvam de assinaturas.

### 5.1 Artigos e trabalhos produzidos

Com os resultados e a técnica implementada, foi possível produzir um artigo intitulado “*Identification of Similar Loads for Electric Power Management in Smart Grid*”. Esse artigo trata da parte inicial do trabalho, em que apenas os resultados com as larguras e número de bandas, escolhidos pontualmente, haviam sido calculados. O artigo foi submetido e aprovado para publicação no IEEE América Latina e atualmente se encontra na fase de revisão com o editor.

Além desse, foi submetido um artigo para o Simpósio Brasileiro de Sistemas Elétricos (SBSE) intitulado “*Intelligent Identification of Similar Loads in Smart Grid*”, que também foi aceito. Entretanto, por se tratar de um trabalho muito semelhante ao submetido na revista supracitada, optou-se pela não participação no evento e consequente não publicação em seus anais.

Um artigo intitulado “Identificação Inteligente de Cargas Elétricas Similares em Smart Grid”, contendo o resultado do trabalho após a implementação do otimizador, também foi aceito para o Congresso Brasileiro de Automatica (CBA). A apresentação do trabalho nesse congresso acontecerá em setembro de 2018.

Por fim, o artigo “Identificação Inteligente de Cargas Elétricas em Smart Grids” foi aceito para a VIII Semana de Produção Científica do Conecta IF, organizado pelo Instituto Federal de Brasília. O trabalho será apresentado em agosto de 2018.

## 5.2 Trabalhos futuros

Para trabalhos futuros, sugere-se:

1. Buscar novas técnicas, tanto de caracterização quanto de classificação, a fim de melhorar a acurácia do sistema de identificação e obter novos resultados comparativos;
2. Reduzir o número de características representativas mantendo a acurácia alcançada;
3. Retornar aos demais classificadores e aplicar o otimizador, de modo a obter a maior taxa de acerto para cada um dos classificadores;
4. Aplicar as técnicas apresentadas em cenários que envolvam cargas com diferentes graus de similaridade, como acontece em casos práticos;
5. Aplicar as técnicas apresentadas em cenários com cargas altamente similares, porém de tipo diferente do que foi tratado neste trabalho. Por exemplo, vários motores com os mesmos dados técnicos e do mesmo fabricante.

# Referências

- ABBAS, G. et al. Solution of an economic dispatch problem through particle swarm optimization: A detailed survey-part i. *IEEE Access*, IEEE, v. 5, p. 15105–15141, 2017. Citado na página 37.
- AHMADI, H.; MARTI, J. R. Load decomposition at smart meters level using eigenloads approach. *IEEE Transactions on Power Systems*, IEEE, v. 30, n. 6, p. 3425–3436, 2015. Citado 3 vezes nas páginas 24, 25 e 28.
- AZZINI, H. A. D.; FARDIN, J. F.; SALLES, E. O. T. Software para monitoramento não intrusivo de cargas residenciais com fins de auditoria energética. *IV Simpósio Brasileiro de Sistemas Elétricos (SBSE)*, 2012. Citado 2 vezes nas páginas 16 e 22.
- BELLEY, C. et al. An efficient and inexpensive method for activity recognition within a smart home based on load signatures of appliances. *Pervasive and Mobile Computing*, Elsevier, v. 12, p. 58–78, 2014. Citado 2 vezes nas páginas 25 e 28.
- BOUHOURLAS, A. S.; MILIOUDIS, A. N.; LABRIDIS, D. P. Development of distinct load signatures for higher efficiency of nilm algorithms. *Electric Power Systems Research*, Elsevier, v. 117, p. 163–171, 2014. Citado 4 vezes nas páginas 19, 28, 29 e 32.
- CALAMERO, N.; BECK, Y.; SHMILOVITZ, D. Defining the unique signatures of loads using the currents' physical components theory and  $z$ -transform. *IEEE Transactions on Industrial Informatics*, IEEE, v. 11, n. 1, p. 155–165, 2015. Citado na página 28.
- CHAKRAVORTY, D.; CHAUDHURI, B.; HUI, S. Estimation of aggregate reserve with point-of-load voltage control. *IEEE Transactions on Smart Grid*, IEEE, 2017. Citado na página 28.
- CHANG, H.-H. et al. A new measurement method for power signatures of nonintrusive demand monitoring and load identification. *IEEE Transactions on Industry Applications*, IEEE, v. 48, n. 2, p. 764–771, 2012. Citado na página 24.
- CHANG, H.-H. et al. Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification. *IEEE Transactions on Industry Applications*, IEEE, v. 50, n. 3, p. 2081–2089, 2014. Citado na página 24.
- CHANG, H.-H. et al. Particle-swarm-optimization-based nonintrusive demand monitoring and load identification in smart meters. *IEEE Transactions on Industry Applications*, IEEE, v. 49, n. 5, p. 2229–2236, 2013. Citado na página 24.
- CHAVES, G. d. L. D.; TOSTA, M. d. C. R. *Gestão de Sistemas de Energia*. [S.l.]: Editora CRV, 2016. 374 p. ISBN 978-85-444-1100-1. Citado na página 14.
- DONG, M. et al. An event window based load monitoring technique for smart meters. *IEEE transactions on smart grid*, IEEE, v. 3, n. 2, p. 787–796, 2012. Citado 2 vezes nas páginas 23 e 24.
- DUAN, J. et al. Power balance control of micro gas turbine generation system based on

- supercapacitor energy storage. *Energy*, Elsevier, v. 119, p. 442–452, 2017. Citado na página 28.
- EHRHARDT-MARTINEZ, K. et al. Advanced metering initiatives and residential feedback programs: a meta-review for household electricity-saving opportunities. In: AMERICAN COUNCIL FOR AN ENERGY-EFFICIENT ECONOMY WASHINGTON, DC. [S.l.], 2010. Citado na página 15.
- ESA, N. F.; ABDULLAH, M. P.; HASSAN, M. Y. A review disaggregation method in non-intrusive appliance load monitoring. *Renewable and Sustainable Energy Reviews*, Elsevier, v. 66, p. 163–173, 2016. Citado na página 28.
- FALCÃO, D. M. Integração de tecnologias para viabilização da smart grid. *III Simpósio Brasileiro de Sistemas Elétricos*, p. 1–5, 2010. Citado 2 vezes nas páginas 14 e 15.
- FAN, Z. et al. Smart grid communications: Overview of research challenges, solutions, and standardization activities. *IEEE Communications Surveys & Tutorials*, IEEE, v. 15, n. 1, p. 21–38, 2013. Citado na página 15.
- FARHANGI, H. The path of the smart grid. *IEEE power and energy magazine*, IEEE, v. 8, n. 1, 2010. Citado na página 14.
- FARINACCIO, L.; ZMEUREANU, R. Using a pattern recognition approach to disaggregate the total electricity consumption in a house into the major end-uses. *Energy and Buildings*, Elsevier, v. 30, n. 3, p. 245–259, 1999. Citado 2 vezes nas páginas 16 e 24.
- FIGUEIREDO, M.; ALMEIDA, A. D.; RIBEIRO, B. Home electrical signal disaggregation for non-intrusive load monitoring (nilm) systems. *Neurocomputing*, Elsevier, v. 96, p. 66–73, 2012. Citado na página 24.
- FIOROTTI, R. et al. Otimização usando pso no controle pi de vazão de um túnel de vento atmosférico utilizando a técnica de gain scheduling. *Latin American Journal of Energy Research*, v. 3, n. 2, p. 30–41, 2017. Citado na página 38.
- GONZALEZ, M.; DEBUSSCHÈRE, V.; BACHA, S. A load identification method for residential building applications. In: IEEE. *Industrial Technology (ICIT), 2012 IEEE International Conference on*. [S.l.], 2012. p. 84–88. Citado na página 28.
- GUL, O.; MILANOVIC, J. Sensitivity of harmonic load model parameters to voltage and current waveforms. In: IEEE. *Harmonics and Quality of Power, 2000. Proceedings. Ninth International Conference on*. [S.l.], 2000. v. 3, p. 1041–1046. Citado na página 28.
- GURSOY, E.; NIEBUR, D. Harmonic load identification using complex independent component analysis. *IEEE Transactions on Power Delivery*, IEEE, v. 24, n. 1, p. 285–292, 2009. Citado na página 28.
- HART, G. W. Nonintrusive appliance load monitoring. *Proceedings of the IEEE*, IEEE, v. 80, n. 12, p. 1870–1891, 1992. Citado na página 16.
- HARTIGAN, J. A.; WONG, M. A. Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, JSTOR, v. 28, n. 1, p. 100–108, 1979. Citado na página 35.

- HASSAN, N. U. et al. Impact of scheduling flexibility on demand profile flatness and user inconvenience in residential smart grid system. *Energies*, Multidisciplinary Digital Publishing Institute, v. 6, n. 12, p. 6608–6635, 2013. Citado na página 28.
- HASSAN, T.; JAVED, F.; ARSHAD, N. An empirical investigation of vi trajectory based load signatures for non-intrusive load monitoring. *IEEE Transactions on Smart Grid*, IEEE, v. 5, n. 2, p. 870–878, 2014. Citado 3 vezes nas páginas 23, 24 e 28.
- HE, D. et al. Front-end electronic circuit topology analysis for model-driven classification and monitoring of appliance loads in smart buildings. *IEEE Transactions on Smart Grid*, IEEE, v. 3, n. 4, p. 2286–2293, 2012. Citado na página 28.
- HELMI, K.; TAYLOR, T.; ANSARI, F. Shear force–based method and application for real-time monitoring of moving vehicle weights on bridges. *Journal of Intelligent Material Systems and Structures*, SAGE Publications Sage UK: London, England, v. 26, n. 5, p. 505–516, 2015. Citado na página 28.
- HSU, Y.-Y.; YANG, C.-C. Design of artificial neural networks for short-term load forecasting. part 1: Self-organising feature maps for day type identification. In: IET. *IEE Proceedings C (Generation, Transmission and Distribution)*. [S.l.], 1991. v. 138, n. 5, p. 407–413. Citado na página 28.
- KENNEDY, J.; EBERHART, R. Pso optimization. In: IEEE SERVICE CENTER, PISCATAWAY, NJ. *Proc. IEEE Int. Conf. Neural Networks*. [S.l.], 1995. v. 4, p. 1941–1948. Citado 2 vezes nas páginas 37 e 38.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: MONTREAL, CANADA. *Ijcai*. [S.l.], 1995. v. 14, n. 2, p. 1137–1145. Citado na página 36.
- KOLODNER, J. *Case-based reasoning*. [S.l.]: Morgan Kaufmann, 2014. Citado na página 35.
- KOUZELIS, K. et al. Estimation of residential heat pump consumption for flexibility market applications. *IEEE Transactions on Smart Grid*, IEEE, v. 6, n. 4, p. 1852–1864, 2015. Citado na página 28.
- LAM, H. Y.; FUNG, G.; LEE, W. A novel method to construct taxonomy electrical appliances based on load signaturesof. *IEEE Transactions on Consumer Electronics*, IEEE, v. 53, n. 2, 2007. Citado na página 28.
- LIANG, J. et al. Load signature study—part i: Basic concept, structure, and methodology. *IEEE transactions on power Delivery*, IEEE, v. 25, n. 2, p. 551–560, 2010. Citado na página 22.
- PAIXÃO, A. R. da. Sistema de classificação inteligente de cargas elétricas. *Dissertação de Mestrado em Energia*, Universidade Federal do Espírito Santo, 2016. Citado 12 vezes nas páginas 17, 18, 19, 20, 23, 26, 30, 32, 35, 36, 40 e 43.
- PAIXÃO, A. R. da et al. Classificação inteligente aplicada ao problema de identificação de cargas elétricas "idênticas". *Simpósio Brasileiro de Pesquisa Operacional*, Anais do XLVIII SBPO, 2016. Citado 4 vezes nas páginas 40, 41, 47 e 48.

- PARK, D. C. et al. Electric load forecasting using an artificial neural network. *IEEE transactions on Power Systems*, IEEE, v. 6, n. 2, p. 442–449, 1991. Citado na página 28.
- RAYUDU, R.; TYLER, C.; WITHERDEN, M. Towards a tool for non-intrusive individual load measurement and identification. In: IEEE. *Innovative Smart Grid Technologies Asia (ISGT), 2011 IEEE PES*. [S.l.], 2011. p. 1–8. Citado na página 28.
- RUZZELLI, A. G. et al. Real-time recognition and profiling of appliances through a single electricity sensor. In: IEEE. *Sensor Mesh and Ad Hoc Communications and Networks (SECON), 2010 7th Annual IEEE Communications Society Conference on*. [S.l.], 2010. p. 1–9. Citado na página 24.
- SADEGHIANPOURHAMAMI, N. et al. Comprehensive feature selection for appliance classification in nilm. *Energy and Buildings*, Elsevier, 2017. Citado 2 vezes nas páginas 25 e 28.
- SINGH, M. et al. Residential load signature analysis for their segregation using wavelet—svm. In: *Power Electronics and Renewable Energy Systems*. [S.l.]: Springer, 2015. p. 863–871. Citado na página 28.
- SRINIVASAN, D.; NG, W.; LIEW, A. Neural-network-based signature recognition for harmonic source identification. *IEEE Transactions on Power Delivery*, IEEE, v. 21, n. 1, p. 398–405, 2006. Citado na página 40.
- SWAMINATHAN, V.; GILBERT, J. L. Fretting corrosion of cochrmo and ti6al4v interfaces. *Biomaterials*, Elsevier, v. 33, n. 22, p. 5487–5503, 2012. Citado na página 28.
- VESTRONI, F.; VIDOLI, S. Closed-form solutions for the structural response to train loads. *Journal of sound and vibration*, Elsevier, v. 303, n. 3, p. 691–706, 2007. Citado na página 28.
- WANG, K. et al. A two-layer framework for quantifying demand response flexibility at bulk supply points. *IEEE Transactions on Smart Grid*, IEEE, 2016. Citado na página 28.
- YALLA, S. K.; KAREEM, A. Dynamic load simulator: Actuation strategies and applications. *Journal of engineering mechanics*, American Society of Civil Engineers, v. 133, n. 8, p. 855–863, 2007. Citado na página 28.
- ZEIFMAN, M.; ROTH, K. Nonintrusive appliance load monitoring: Review and outlook. *IEEE transactions on Consumer Electronics*, IEEE, v. 57, n. 1, 2011. Citado 2 vezes nas páginas 19 e 22.