

UNIVERSIDADE DO ESTADO DE SANTA CATARINA – UDESC
CENTRO DE CIÊNCIAS TECNOLÓGICAS – CCT
DEPARTAMENTO DE ENGENHARIA ELÉTRICA – DEE

VITOR LEITE DE BARROS HEINZLE

**APLICAÇÃO DA TÉCNICA NILM PARA DESAGREGAÇÃO DE CARGAS
RESIDENCIAIS UTILIZANDO MÉTODO BASEADO EM EVENTOS**

JOINVILLE

2023

VITOR LEITE DE BARROS HEINZLE

**APLICAÇÃO DA TÉCNICA NILM PARA DESAGREGAÇÃO DE CARGAS
RESIDENCIAIS UTILIZANDO MÉTODO BASEADO EM EVENTOS**

Trabalho de Conclusão de Curso
apresentado ao curso de Bacharelado em
Engenharia Elétrica do Centro de Ciências
Tecnológicas da Universidade do Estado
de Santa Catarina, como requisito para a
obtenção do título de Bacharel em
Engenharia Elétrica.

Orientador: Prof. Dr. Sérgio Vidal Garcia
Oliveira

JOINVILLE

2023

VITOR LEITE DE BARROS HEINZLE

**APLICAÇÃO DA TÉCNICA NILM PARA DESAGREGAÇÃO DE CARGAS
RESIDENCIAIS UTILIZANDO MÉTODO BASEADO EM EVENTOS**

Trabalho de Conclusão de Curso
apresentado ao curso de Bacharelado em
Engenharia Elétrica do Centro de Ciências
Tecnológicas da Universidade do Estado
de Santa Catarina, como requisito para a
obtenção do título de Bacharel em
Engenharia Elétrica.

Orientador: Prof. Dr. Sérgio Vidal Garcia
Oliveira

BANCA EXAMINADORA

Prof. Dr. Sérgio Vidal Garcia Oliveira
Universidade do Estado de Santa Catarina

Membros:

Prof. Dr. Alessandro Luiz Batschauer
Universidade do Estado de Santa Catarina

Prof. Dr. Fabiano Ferreira Andrade
Universidade do Estado de Santa Catarina

Joinville, 15 de dezembro de 2023

“Prometo que, no cumprimento do meu dever de engenheiro, não me deixarei cegar pelo brilho excessivo da tecnologia, não me esquecendo de que trabalho para o bem do homem, e não da máquina. Respeitarei a natureza, evitando projetar ou construir equipamentos que destruam o equilíbrio ecológico ou poluam. Colocarei todo o meu conhecimento científico a serviço do conforto e desenvolvimento da humanidade.

Assim estarei em paz comigo e com Deus.

Assim eu juro.”

Juramento do Curso de Engenharia Elétrica
Universidade do Estado de Santa Catarina

À minha Mãe.

Eu tenho tanto pra lhe falar
Mas com palavras não sei dizer
Como é grande o meu amor por você
E não há nada pra comparar
Para poder lhe explicar
Como é grande o meu amor por você
Nem mesmo o céu nem as estrelas
Nem mesmo o mar e o infinito
Nada é maior que o meu amor
Nem mais bonito
Me desespero a procurar
Alguma forma de lhe falar
Como é grande o meu amor por você
Nunca se esqueça nem um segundo
Que eu tenho o amor maior do mundo
Como é grande o meu amor por você
Nunca se esqueça, nem um segundo
Que eu tenho o amor maior do mundo
Como é grande o meu amor por você
Mas como é grande o meu amor por você

Como É Grande O Meu Amor Por Você
(Roberto Carlos)

AGRADECIMENTOS

À Universidade do Estado de Santa Catarina, pela oportunidade do aprendizado, em especial:

Ao meu orientador, por ter acreditado em mim. Os desafios que plantou em meu caminho floresceram e contribuíram não somente para este trabalho, mas para meu crescimento profissional e amadurecimento pessoal. Seu profissionalismo me inspira a ser um melhor engenheiro.

Aos professores Alessandro Batschauer e Fabiano Andrade, por aceitarem o convite de participar da banca. Suas contribuições foram valiosas para o aprimoramento deste trabalho.

Aos professores do Departamento de Engenharia Elétrica. Seus ensinamentos trarão prosperidade a mim e a futuras gerações.

Agradeço também:

A meus pais, Alvaro e Ângela, que sempre fizeram de tudo por mim. Este trabalho é dedicado a vocês.

A minha família, em especial meus avós Cecília Villaça e Ralf Heinzle (*In Memoriam*), que me inspiram a lutar e dedicar-me.

Aos meus amigos que me acompanham nesta caminhada.

A Deus que ilumina e capacita a todos nós.

A todos, o meu muito obrigado!

RESUMO

O monitoramento do consumo energético de cada aparelho em uma instalação elétrica permite identificar desperdícios e promover a eficiência energética. A técnica tradicional de monitoramento emprega sensores de consumo instalados em cada carga. Embora essa abordagem ofereça alta precisão nas medições, resulta em custos elevados. A técnica de Monitoramento de Cargas Não Intrusivo (NILM) surge como uma solução inovadora para superar as limitações da técnica tradicional e possibilitar a análise do consumo energético de cada carga a partir de um único ponto de medição. No cerne da técnica NILM estão as assinaturas elétricas, autênticas "impressões digitais" das cargas, distinguíveis por características mensuráveis como potência ativa e reativa. A técnica opera através de um sistema de aquisição que coleta a potência total consumida (potência agregada). Enquanto isso, algoritmos especializados detectam variações no sinal adquirido, denominadas eventos, correlacionando-os com as assinaturas conhecidas para desagregar o consumo e identificar as cargas ativas. Este estudo tem como foco explorar o algoritmo de Otimização Combinatória (CO), utilizando o software NILMTK, para avaliar seu desempenho na identificação de cargas em uma residência da base de dados REDD, contribuindo para aprimorar a gestão energética residencial.

Palavras-chave: Assinaturas Elétricas; Cargas Residenciais; Eficiência Energética; Monitoramento de Cargas Não Intrusivo; Otimização Combinatória.

ABSTRACT

Monitoring the energy consumption of each appliance in an electrical installation allows for the identification of wastage and promotes energy efficiency. The traditional monitoring technique employs consumption sensors installed on each load. Although this approach offers high measurement precision, it results in elevated costs. The Non-Intrusive Load Monitoring (NILM) technique emerges as an innovative solution to overcome the limitations of the traditional technique, enabling the analysis of the energy consumption of each load from a single measurement point. At the heart of the NILM technique are the electrical signatures, authentic "fingerprints" of the loads, distinguishable by measurable characteristics such as active and reactive power. The technique operates through an acquisition system that collects the total power consumed (aggregated power). Meanwhile, specialized algorithms detect variations in the acquired signal, termed events, correlating them with known signatures to disaggregate the consumption and identify the active loads. This study focuses on exploring the Combinatorial Optimization (CO) algorithm, utilizing the NILMTK software, to assess its performance in identifying loads in a residence from the REDD database, contributing to the improvement of residential energy management.

Keywords: Electrical Signatures; Residential Loads; Energy Efficiency; Non-Intrusive Load Monitoring; Combinatorial Optimization.

LISTA DE ILUSTRAÇÕES

Figura 1 — Ilustração de um sistema de aquisição de potência agregada.	20
Figura 2 — Exemplo de desagregação da potência ativa de uma residência.	21
Figura 3 — Ilustração dos modelos de tipos de cargas elétricas proposta por (HART, 1992). (a) Tipo Liga/Desliga, (b) Máquina de Estados Finitos, (c) e (d) Continuamente Variável.	29
Figura 4 — Ilustração conceitual de eventos de subida e descida e estados transitórios e em regime permanente.	36
Figura 5 — Fluxograma do Algoritmo CO, correlacionando: as três etapas de um sistema NILM (a), fase de treinamento (b) e fase de desagregação (c).	41
Figura 6 — Representação do modelo de assinaturas elétricas formado por <i>clusters</i> de potências ativa e reativa.	42
Figura 7 — Consumo de potência ativa da geladeira da Residência 1 ao longo de três dias.	51
Figura 8 — Potência agregada medida na Residência 1 entre os dias 20/04 e 27/04/2011.	52
Figura 9 — Consumo de potência ativa da geladeira da Residência 1, com detalhe para o momento de acionamento do seu sistema de degelo.	53
Figura 10 — Consumo de potência ativa ao longo de um ciclo completo da Lavadora e secadora de Roupas.	54
Figura 11 — Consumo de potência ativa do circuito de iluminação (1ª Instância). ...	55
Figura 12 — Consumo de potência ativa do circuito de tomadas (2ª Instância) ao longo de uma semana (20/04 a 27/04/2011).	56
Figura 13 — Consumo de potência ativa da Lava-Louças.	56
Figura 14 — Consumo de potência ativa do Micro-ondas.	58
Figura 15 — F1-Score do cenário de desagregação das TopK 10 cargas com a janela de treinamento de dois dias.	64
Figura 16 — Resultado da desagregação da geladeira durante aproximadamente um dia (1º Cenário).	65
Figura 17 — Resultado da desagregação da geladeira no mesmo período mostrado na Figura 16.	65
Figura 18 — Detalhe do momento de acionamento do compressor da geladeira, com sobreposição do consumo individual e agregado cargas.	66

Figura 19 — Detalhe do resultado da desagregação em um ciclo de acionamento da geladeira, evidenciando erros no processo (1º Cenário).	66
Figura 20 — F1-Score da desagregação das TopK 13 cargas com a janela de treinamento de dois dias.	67
Figura 21 — F-Score do cenário de desagregação de todas as cargas instaladas na residência com a janela de treinamento de dois dias.	68
Figura 22 — Detalhe do resultado da desagregação em um dos ciclos de acionamento da geladeira (3º Cenário).	68
Figura 23 — F1-Score do cenário de desagregação das TopK 10 cargas com a janela de treinamento de um dia.	69
Figura 24 — F1-Score do cenário de desagregação das TopK 13 cargas com a janela de treinamento de um dia.	70
Figura 25 — Comparação da métrica F1-Score para todos os cenários testados. ...	71

LISTA DE ABREVIATURAS E SIGLAS

BLUED	<i>Building-level fully-labeled dataset for electricity disaggregation</i>
CO	Otimização Combinatória — <i>Combinatorial Optimization</i>
IDE	Interface de Desenvolvimento Integrado
ILM	Monitoramento de Cargas Intrusivo — <i>Intrusive Load Monitoring</i>
MIT	Instituto de Tecnologia de Massachusetts — <i>Massachusetts Institute of Technology</i>
NILM	Monitoramento de Cargas Não Intrusivo — <i>Nonintrusive Load Monitoring</i>
PCC	Ponto de Conexão Comum
REDD	<i>Reference Energy Disaggregation Data Set</i>
SGE	Sistemas de Gestão de Energia
SCP	Princípio da Unicidade de Eventos — <i>Switch Continuity Principle</i>
UK-DALE	<i>United Kingdom Domestic Appliance Level Electricity</i>

LISTA DE VARIÁVEIS

$a_i(t)$	Vetor booleano de tamanho n , que indica o estado de cada uma das n cargas em um instante t .
$\hat{a}(t)$	Vetor booleano representando o conjunto de estados a_i das cargas que resultam no mínimo erro ou discrepância entre o consumo agregado medido no instante t , $P_{Total}(t)$, e o somatório dos consumos estimados para todos os aparelhos individuais.
$e(t)$	Qualquer contribuição em potência de aparelhos não considerados mais o ruído medido.
i	Índice usado para enumerar as cargas de 1 a n .
$P_{Total}(t)$	Potência agregada.
P_i	Potência consumida por cada i aparelho.
$P_{Norm}(t)$	Potência corrigida.
V_{Nom}	Tensão nominal de fornecimento da rede.
$V(t)$	Tensão medida ao longo do tempo.
$Y(t)$	Admitância total observada a partir do ponto medido.
TP	Número de vezes que uma classe de eventos foi corretamente identificada (sucesso de classificação).
FP	Número de vezes que uma classe de eventos foi prevista sem que ela tenha ocorrido (erro de classificação).
FN	Número de vezes que uma classe não foi prevista pelo algoritmo (erro de classificação).

SUMÁRIO

1	INTRODUÇÃO	15
1.1	OBJETIVOS	17
1.1.1	Objetivo Geral	17
1.1.2	Objetivos Específicos	17
2	DESENVOLVIMENTO TEÓRICO	19
2.1	INTRODUÇÃO À TÉCNICA NILM.....	19
2.1.1	Motivação e Contexto	19
2.1.2	Princípios de Funcionamento	21
2.1.3	Desafios Técnicos	22
2.1.4	Classificações dos Sistemas NILM.....	23
2.1.5	Aplicações da Técnica	24
2.1.5.1	<i>Aplicações para Consumidores</i>	<i>24</i>
2.1.5.2	<i>Aplicações para Fornecedores de Produtos e Serviços</i>	<i>26</i>
2.2	CARACTERIZAÇÃO DO PROBLEMA	27
2.3	TIPOS DE CARGAS.....	29
2.3.1	Tipo 1: Cargas Liga/Desliga	29
2.3.2	Tipo 2: Cargas com Estados Finitos.....	30
2.3.3	Tipo 3: Cargas com Variação Contínua.....	30
2.3.4	Tipo 4: Cargas Permanentes	30
2.4	ASSINATURAS ELÉTRICAS	30
2.4.1	Regime Permanente	31
2.4.2	Transitórios.....	33
2.5	AQUISIÇÃO DOS SINAIS	33
2.5.1	Frequência de Aquisição	33
2.5.2	Resolução e Precisão dos Dados	34
2.6	EXTRAÇÃO DE PROPRIEDADES	35
2.6.1	Métodos Baseados em Eventos.....	36
2.6.2	Métodos Baseados em Estados.....	37
2.7	ALGORITMOS DE CLASSIFICAÇÃO	37
2.7.1	Supervisionados.....	38
2.7.2	Não Supervisionados.....	38
2.7.3	O Algoritmo CO	39

2.7.3.1	Treinamento	40
2.7.3.2	Desagregação	42
2.8	SOFTWARE NILMTK	44
2.9	MÉTRICAS DE DESEMPENHO.....	45
2.9.1	Precisão	45
2.9.2	Recall.....	46
2.9.3	F1-Score	46
2.10	BASES DE DADOS.....	46
2.11	CONCLUSÃO.....	47
3	METODOLOGIA PROPOSTA	49
3.1	DEFINIÇÃO DO CONJUNTO DE DADOS UTILIZADOS	49
3.1.1	Banco de Dados	49
3.1.2	A Residência Selecionada e Suas Cargas Relevantes.....	50
3.1.2.1	Geladeira.....	53
3.1.2.2	Lavadora e Secadora de Roupas.....	53
3.1.2.3	Circuito de Iluminação	54
3.1.2.4	Circuito de Tomadas (2ª Instância)	55
3.1.2.5	Lava-louças	55
3.1.2.6	Micro-ondas.....	56
3.2	CENÁRIOS DE INVESTIGAÇÃO PROPOSTOS	58
3.2.1	Cenário 1: 2D e TopK = 10	59
3.2.2	Cenário 2: 2D e TopK = 13	59
3.2.3	Cenário 3: 2D e Todas as Cargas da Residência (TopK = 16)	59
3.2.4	Cenário 4: 1D e TopK = 10	60
3.2.5	Cenário 5: 1D e TopK = 13	60
3.3	TREINAMENTO DO ALGORITMO.....	60
3.4	DESAGREGAÇÃO	61
3.5	MÉTODOS DE AVALIAÇÃO DOS RESULTADOS	61
3.5.1	Avaliação Qualitativa: Uso de Gráficos.....	61
3.5.2	Avaliação Quantitativa: F1-Score.....	61
3.6	CONCLUSÃO.....	62
4	RESULTADOS.....	63
4.1	DESAGREGAÇÃO DAS CARGAS DE MAIOR CONSUMO	63
4.1.1	Janela de Treinamento de Dois Dias (2D)	63

4.1.1.1	<i>Cenário 1: 2D e TopK = 10.....</i>	<i>64</i>
4.1.1.2	<i>Cenário 2: 2D e TopK = 13.....</i>	<i>67</i>
4.1.1.3	<i>Cenário 3: 2D e Todas as Cargas da Residência (TopK 16).....</i>	<i>67</i>
4.1.2	Janela de Treinamento de Um Dia	69
4.1.2.1	<i>Cenário 4: 1D e TopK = 10.....</i>	<i>69</i>
4.1.2.2	<i>Cenário 5: 1D e TopK = 13.....</i>	<i>70</i>
4.2	COMPARAÇÃO DOS RESULTADOS.....	70
4.3	DISCUSSÃO	71
4.3.1	Consolidação do Aprendizado com Janelas de Treinamento Reduzidas	71
4.3.2	Desempenho Comprometido por Assinaturas Elétricas Similares....	72
4.3.3	Acréscimo de Cargas Conhecidas e Melhoria de Desempenho	72
4.3.4	Análise Específica: A Geladeira.....	73
4.3.5	Observações Sobre o Funcionamento do Algoritmo	73
4.3.6	Uma Breve Discussão dos Impactos do Erro $e(t)$.....	74
4.3.7	Aquecedor de Ambientes e Cargas Desconhecidas	74
5	CONCLUSÕES, SUGESTÕES DE TRABALHOS FUTUROS E CONSIDERAÇÕES FINAIS	75
5.1	SUGESTÕES DE TRABALHOS FUTUROS	77
5.2	CONSIDERAÇÕES FINAIS	77
	REFERÊNCIAS.....	79

1 INTRODUÇÃO

As mudanças climáticas e seus impactos representam desafios globais sem precedentes, exigindo ação imediata e decisiva para mitigar seus efeitos nocivos. A matriz energética mundial, intimamente ligada às emissões de gases de efeito estufa, desempenha um papel determinante nesse cenário, tornando-se imperativo a adoção de práticas de consumo responsável e eficiente de energia. No Brasil, especificamente em 2022, o setor residencial representou 30% do consumo de eletricidade do país (BRASIL, 2022), com eletrodomésticos como refrigeradores, TVs, chuveiros, ar-condicionado, lâmpadas, ventiladores, freezers, micro-ondas e máquinas de lavar roupa liderando este consumo (BRASIL, 2019). Estes aparelhos apresentam características elétricas distintas, cujos padrões de consumo podem ser cuidadosamente identificados.

A análise detalhada destes padrões de consumo revela uma oportunidade significativa para a implementação de estratégias direcionadas à eficiência energética. O extenso *review* da literatura feito por (CARRIE ARMEL, GUPTA, *et al.*, 2013) revela que o fornecimento de *feedback* detalhado sobre o consumo individual dos aparelhos pode influenciar positivamente os hábitos dos consumidores, levando a uma redução média de 12% no consumo de eletricidade residencial nos estudos avaliados. Adicionalmente, o autor destaca que a disponibilização deste *feedback* em tempo real, combinado a sua integração a sistemas de gestão de cargas automatizado, amplia ainda mais o potencial de eficiência energética.

É neste contexto que a técnica *Nonintrusive Load Monitoring* (NILM — Monitoramento de Cargas Não Intrusivo) destaca-se como uma abordagem promissora, permitindo uma gestão energética eficiente em ambientes residenciais, comerciais e industriais por meio da identificação de aparelhos conectados à rede elétrica sem a necessidade de instalação de sensores dedicados para cada carga. Contrapondo-se ao *Intrusive Load Monitoring* (ILM — Monitoramento de Cargas Intrusivo), que requer a instalação direta de sensores em cada dispositivo para monitoramento, a NILM opera pela análise de sinais agregados, que representam o consumo total de energia de uma instalação. Esta técnica aplica engenharia reversa nos sinais de potência agregada utilizando algoritmos avançados de processamento para identificar os padrões de consumo específicos de cada aparelho, marcando um avanço significativo sobre a abordagem mais invasiva e menos prática da ILM.

No cerne da técnica NILM estão as assinaturas elétricas, autênticas "impressões digitais" das cargas, distinguíveis por características mensuráveis como potência ativa, reativa, frequência de operação e padrões de uso sazonais. Essas assinaturas possibilitam que o sistema NILM realize a identificação e o monitoramento do consumo energético de cada carga de forma individualizada. A metodologia NILM é caracterizada pela diversidade de abordagens, abrangendo uma variedade de algoritmos, frequências de aquisição e tipos de assinaturas elétricas utilizadas para a identificação, o que permite a sua adaptação conforme a aplicação desejada.

Motivado pela oportunidade de analisar em detalhes algumas cargas elétricas residenciais e pela possibilidade de descobrir oportunidades de eficiência energética em um contexto familiar, este estudo se propõe a investigar o primeiro algoritmo de desagregação NILM baseado no método de Otimização Combinatória (CO), proposto por (HART, 1992). Este marco inicial na área introduziu conceitos fundamentais que continuam sendo aplicados em técnicas contemporâneas. A exploração desses conceitos, juntamente com a aplicação do algoritmo CO, facilita o entendimento dos princípios subjacentes à metodologia NILM, demonstrando sua aplicabilidade e relevância duradoura.

Este trabalho divide-se nas seguintes seções.

Capítulo 1: Introdução e Objetivos — Esta seção introduz o tema abordado e define os objetivos gerais e específicos do estudo, estabelecendo a fundação para as análises subsequentes.

Capítulo 2: Revisão Bibliográfica — Oferece uma análise detalhada dos conceitos chave associados à técnica NILM, iniciando com uma contextualização da técnica, suas vantagens, desafios e as aplicações destacadas na literatura. Prossegue com a caracterização do problema de desagregação de potência, os diferentes tipos de cargas e a importância das assinaturas elétricas na identificação das cargas. Em seguida, detalha as três etapas fundamentais dos sistemas NILM (Aquisição dos Sinais, Extração de Propriedades e Classificação das Cargas), focando na técnica específica explorada neste trabalho. Finaliza apresentando as ferramentas utilizadas, incluindo o software NILMTK, a métrica de desempenho F-Score, além de mencionar alguns bancos de dados relevantes.

Capítulo 3: Metodologia Proposta — Descreve com precisão a metodologia adotada no estudo, incluindo a seleção e o tratamento dos dados. Especifica os critérios para escolha do ambiente residencial estudado e os cenários investigados,

visando avaliar a eficácia do algoritmo CO, com orientações claras para futura replicação ou expansão da pesquisa.

Capítulo 4: Resultados — Apresenta os resultados alcançados mediante a aplicação prática da metodologia proposta, analisando o desempenho do algoritmo em diferentes cargas e cenários. Inclui uma discussão qualitativa e quantitativa sobre os achados, refletindo sobre como os conceitos técnicos previamente abordados impactam no desempenho do sistema, destacando pontos fortes e limitações identificadas.

Capítulo 5: Conclusões, Sugestões de Trabalhos Futuros e Considerações Finais: Sumariza os achados principais do estudo, realçando a importância do método CO para a prática do NILM e seu potencial para diversas aplicações. Sugere direções para investigações futuras e fornece considerações finais acerca da pesquisa.

Referências: Lista todas as referências bibliográficas consultadas e citadas ao longo do trabalho, fundamentais para a sustentação teórica e metodológica do estudo.

1.1 OBJETIVOS

A seguir são descritos os objetivos gerais e específicos que norteiam este trabalho.

1.1.1 *Objetivo Geral*

O objetivo geral deste estudo consiste em investigar experimentalmente o funcionamento do algoritmo CO utilizando sua implementação no *software* NILMTK, visando avaliar seu desempenho na desagregação de cargas elétricas.

1.1.2 *Objetivos Específicos*

Os objetivos específicos são:

- Definir um banco de dados adequado capaz de fornecer dados de potência ativa agregada e individual das cargas de uma residência real;

- Definir uma métrica de desempenho para o algoritmo CO, possibilitando sua avaliação quantitativa;
- Analisar as assinaturas elétricas das cargas e classificá-las quanto ao seu tipo e perfil de utilização;
- Investigar a influência do tempo de treinamento e do número de cargas conhecidas no desempenho do algoritmo CO, variando tais parâmetros em diferentes cenários experimentais;
- Executar o processo de treinamento e desagregação utilizando o módulo *CombinatorialOptimization* do software NILMTK empregando os dados da residência escolhida;
- Avaliar os resultados da desagregação utilizando módulo de métricas do software NILMTK;
- Estabelecer correlações entre os resultados da desagregação, as assinaturas elétricas e os cenários propostos, buscando inferir conclusões aplicáveis a outros contextos e identificar os principais fatores que influenciam o desempenho do algoritmo CO e;
- Identificar cenários reais de consumo de energia que possam se beneficiar da metodologia NILM baseada em eventos.

2 DESENVOLVIMENTO TEÓRICO

Neste capítulo, são abordados os conceitos teóricos fundamentais que fornecem a base para a compreensão do trabalho realizado nesta pesquisa. A maioria desses fundamentos foi pioneiramente proposta pelo pesquisador George W. Hart, em sua obra *Nonintrusive Appliance Load Monitoring* (1992). Em sua obra, Hart apresenta um estudo abrangente sobre os métodos NILM, abordando desde os princípios básicos até aplicações práticas. As pesquisas de Hart foram fundamentais para o desenvolvimento desta área de pesquisa e seus conceitos continuam a ser relevantes para os estudos contemporâneos.

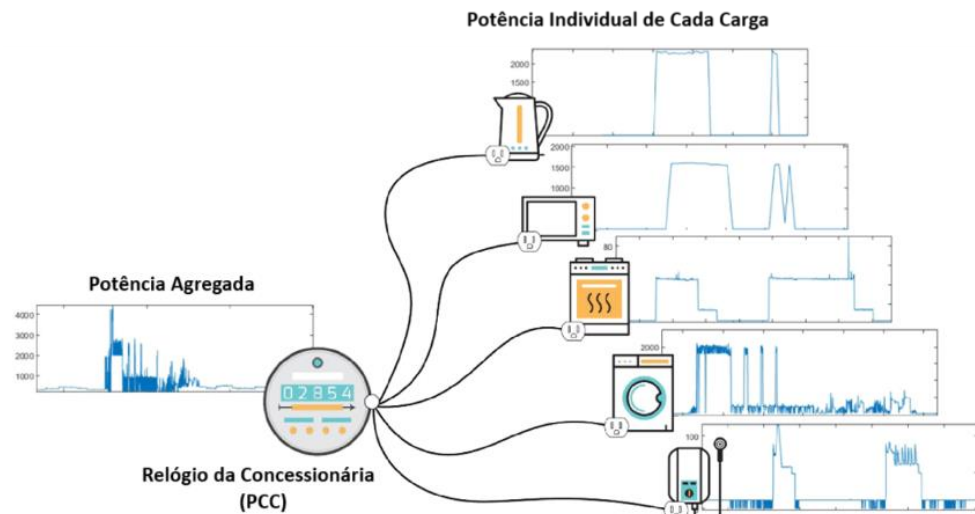
2.1 INTRODUÇÃO À TÉCNICA NILM

Esta seção dedica-se a introduzir conceitos da técnica partindo de uma contextualização ampla para detalhes mais específicos do seu funcionamento e aplicações.

2.1.1 *Motivação e Contexto*

O monitoramento do consumo energético de cada aparelho em uma residência ou instalação elétrica permite identificar desperdícios e promover a eficiência energética. A técnica tradicional de monitoramento, denominada ILM, emprega sensores de consumo instalados em cada carga. Embora essa abordagem ofereça alta precisão nas medições, resulta em custos elevados e complexidade operacional devido ao grande número de sensores necessários, os quais devem ser distribuídos por toda a área de interesse. Nesse contexto, a técnica NILM surge como uma solução inovadora para superar as limitações do ILM ao possibilitar a análise do consumo energético de cada carga a partir de um único ponto de medição, conhecido como Ponto de Conexão Comum (PCC), geralmente localizado na interface entre as instalações do consumidor e da concessionária. A Figura 1 ilustra uma instalação elétrica onde todas as cargas estão conectadas ao PCC. Seus sinais de consumo individuais se sobrepõem e dão origem ao sinal de potência agregada.

Figura 1 — Ilustração de um sistema de aquisição de potência agregada.



Fonte: Adaptado de (PUJIĆ, JELIĆ, *et al.*, 2020).

As vantagens dessa abordagem sobre o ILM incluem (FAUSTINE, MVUNGI, *et al.*, 2017):

- Custo reduzido: a técnica necessita de um número significativamente menor de sensores, tornando a instalação mais econômica e viável;
- Menor invasão: a instalação do NILM é simplificada e não requer modificações na estrutura da instalação elétrica, tornando-a ideal para residências e ambientes já existentes;
- Escalabilidade: a técnica permite monitorar um vasto número de cargas sem que isso implique um aumento proporcional no número de sensores.

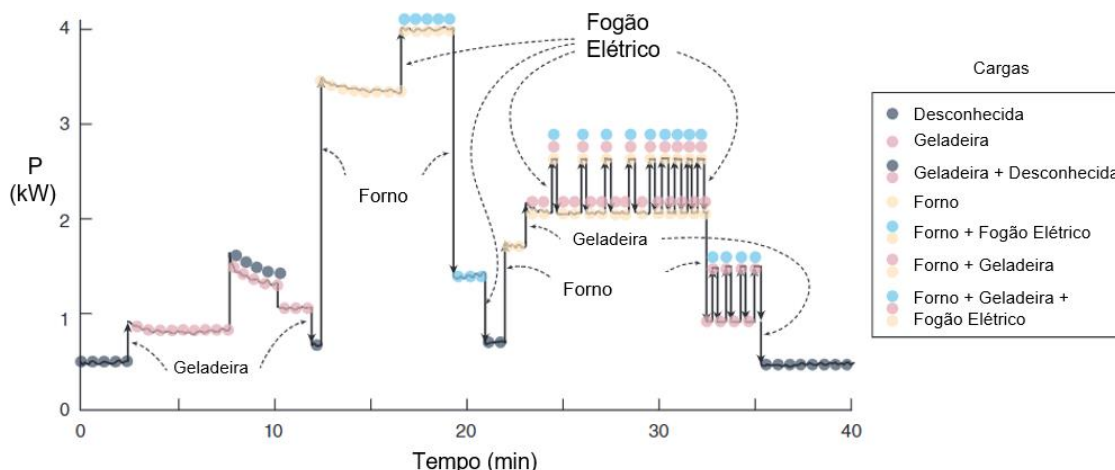
Apesar das vantagens, a técnica NILM também enfrenta alguns desafios técnicos que podem comprometer sua precisão, tornando-a, em certas circunstâncias, menos eficaz que a abordagem ILM. Tais desafios incluem a precisão dos algoritmos de desagregação, a sensibilidade a interferências e ruídos, bem como a exigência de uma base de dados de treinamento abrangente para os aparelhos. Para discutir adequadamente essas limitações, é essencial primeiro entender os princípios de funcionamento do NILM. Assim, a próxima subseção detalha esses fundamentos, preparando o leitor para uma análise dos desafios técnicos.

2.1.2 Princípios de Funcionamento

A técnica NILM pode ser dividida em três etapas sequenciais: Aquisição dos Sinais, Extração de Propriedades e Classificação das Cargas (ABUBAKAR, KHALID, *et al.*, 2015). Na primeira etapa, um sistema de medição localizado no PCC coleta os sinais de tensão e corrente e calcula a potência total consumida pelas cargas, denominada potência agregada. Esta representa a soma das potências consumidas por todas as cargas na instalação. A segunda etapa consiste em identificar variações, denominadas eventos, nos níveis de potência agregada, que estão associados aos ciclos de operação das cargas, como o ato de ligar e desligar uma carga ou durante a transição entre ciclos de operação em aparelhos como máquinas de lavar. Considerando que a maioria das cargas residenciais opera segundo padrões previsíveis, os sistemas NILM se baseiam neste princípio para relacionar cada padrão de consumo, denominado assinatura elétrica, a uma carga específica. A etapa final, Classificação das Cargas, utiliza algoritmos de aprendizado de máquina para efetuar uma engenharia reversa do sinal agregado. Por meio da análise e correlação dos eventos detectados com as assinaturas elétricas previamente conhecidas, o sistema é capaz de desagregar o sinal e identificar as cargas ativas, alcançando o objetivo principal da técnica NILM.

A Figura 2 demonstra a técnica de desagregação de um sinal de potência ativa, identificando o consumo de cada aparelho — geladeira, forno e fogão elétrico — ao longo do tempo e destacando os períodos de uso.

Figura 2 — Exemplo de desagregação da potência ativa de uma residência.



Fonte: Adaptado de (PEREIRA, NUNES, 2018) e (HART, 1992).

Cada uma das etapas da técnica NILM é detalhadamente explorada nas seções 2.5 (Aquisição dos Sinais), 2.6 (Extração de Propriedades) e 2.7 (Classificação das Cargas). A próxima subseção discute os desafios técnicos enfrentados na implementação de um sistema NILM.

2.1.3 Desafios Técnicos

Embora a técnica NILM apresente um grande potencial para o monitoramento energético, sua implementação enfrenta diversos desafios. Isso se deve especialmente à presença de múltiplas cargas com assinaturas elétricas variadas operando simultaneamente, o que complica a desagregação do sinal e a torna uma tarefa complexa. Definir um conjunto de assinaturas elétricas que modele adequadamente cada aparelho, evitando confusões entre as cargas, é crucial e tem implicações em todas as etapas de um sistema NILM, como mostra o estudo de (SADEGHIANPOURHAMAMI, RUYSSINCK, *et al.*, 2017).

A aquisição dos sinais pode ser comprometida por interferências e ruídos tanto no sinal agregado quanto nos decorrentes do sistema de aquisição. A habilidade de filtrar esses ruídos e extrair informações úteis dos dados coletados é fundamental para o sucesso do sistema, conforme destacado por (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022). Além disso, a detecção e classificação de cargas menores representam um desafio significativo, pois suas assinaturas elétricas podem ser ofuscadas pelo ruído ou pela presença de cargas maiores, como discutido em (ZOHA, GLUHAK, *et al.*, 2012). Uma alta frequência e resolução de amostragem podem aprimorar a qualidade das informações disponíveis para análise, mas resultam em um grande volume de dados, exigindo maior capacidade de processamento e armazenamento, o que pode ser inviável para soluções em tempo real ou em dispositivos com recursos limitados, conforme observado por (GOPINATH, KUMAR, *et al.*, 2020).

Na extração de características, determinar as assinaturas elétricas que melhor definem as cargas a partir do sinal amostrado é um desafio. Esta fase envolve a seleção de técnicas de processamento de sinal capazes de lidar com a variedade de eventos e a intermitência dos sinais das cargas. A crescente utilização de cargas acionadas por conversores de frequência e dispositivos eletrônicos inteligentes, que modulam o consumo de energia de formas complexas e variáveis, representa um grande desafio para a técnica NILM (ZOHA, GLUHAK, *et al.*, 2012). A performance e

a interpretação dos resultados podem variar consideravelmente dependendo de como essa fase é executada, segundo (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022).

O principal desafio na etapa de classificação das cargas é a escolha do algoritmo de desagregação, dado que várias técnicas existem, mas nenhuma solução universal se aplica a todos os tipos de cargas e cenários (REHMAN, LIE, *et al.*, 2021). A adaptabilidade dos algoritmos de desagregação a mudanças ao longo do tempo é crítica. A introdução de novos aparelhos, retirada de outros e variações nos padrões de uso exigem soluções NILM robustas e adaptáveis. Nesse contexto, a construção de um banco de dados abrangente, atuando como uma biblioteca de assinaturas elétricas de todas as cargas envolvidas, representa um desafio significativo (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022).

Ao idealizar a técnica NILM há três décadas, (HART, 1992) já antecipava que o sistema NILM ideal não dependeria do fornecimento de assinaturas elétricas coletadas de forma intrusiva, mas seria capaz de aprendê-las diretamente a partir dos sinais agregados. Essa abordagem tem o potencial de dispensar a fase de treinamento do algoritmo, economizando tempo e custos e ampliando a aplicabilidade da solução para cargas não previamente conhecidas. Contudo, essa "solução ideal" ainda não foi totalmente realizada e os sistemas que operam com uma fase de treinamento tendem a apresentar desempenho superior àqueles que aprendem as assinaturas elétricas apenas a partir de dados agregados (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022).

Os desafios mencionados nesta seção ressaltam a necessidade de abordagens robustas e flexíveis na implementação de sistemas NILM. Este trabalho discute essas e outras limitações da técnica, bem como seus impactos no desempenho dos experimentos propostos. A introdução prossegue na subseção seguinte, que aborda as abordagens supervisionadas e não supervisionadas, além de outras classificações relevantes para sistemas NILM.

2.1.4 Classificações dos Sistemas NILM

Os sistemas NILM podem ser divididos em duas categorias principais: supervisionados e não supervisionados. Os sistemas supervisionados requerem uma fase de treinamento prévio, na qual o sistema aprende as características únicas de consumo de cada carga presente na instalação. Por outro lado, os sistemas não

supervisionados operam sem essa fase de treinamento, utilizando algoritmos que buscam identificar padrões de consumo diretamente no sinal de potência agregada. Estes conceitos são detalhados na Seção 2.7.

A diferenciação dos sistemas NILM também ocorre pelo critério de identificação das cargas elétricas e suas assinaturas. O autor (HART, 1992) descreve duas abordagens principais: a baseada em eventos, que detecta mudanças significativas no consumo de energia para identificar diferentes aparelhos sendo ligados ou desligados, e a baseada em estados, que analisa o estado operacional constante dos aparelhos para reconhecê-los. Ambas as estratégias são vantajosas para a análise do consumo de energia em uma instalação e seus conceitos são detalhados na Seção 2.6.

Além disso, os sistemas NILM podem ser distinguidos pelo tipo de assinaturas elétricas utilizadas, dividindo-se em abordagens de regime permanente e regime transitório. Essas categorias, intimamente relacionadas à frequência de aquisição, permitem classificar os sistemas como de baixa ou alta frequência. Estas classificações são exploradas com mais detalhes nas seções 2.4 e 2.5, respectivamente.

A seguir, concluindo a introdução aos sistemas NILM, destacam-se algumas aplicações amplamente discutidas na literatura.

2.1.5 Aplicações da Técnica

A técnica NILM surge como uma ferramenta poderosa para a gestão eficiente de energia, abrindo um leque de aplicações promissoras em diferentes setores, tanto do lado da demanda (consumidores) quanto da oferta (concessionárias e distribuidoras). Na sequência, são apresentadas as principais aplicações da técnica identificadas na literatura.

2.1.5.1 Aplicações para Consumidores

Uma revisão das recentes aplicações da tecnologia NILM voltada para os consumidores de energia é apresentada em (BUCCI, CIANCETTA, et al., 2021), que destaca:

- **Gestão de Energia para Residências Inteligentes:** Esta aplicação se baseia no uso de Sistemas de Gestão de Energia (SGE) que, integrados à infraestrutura de casas inteligentes, oferecem uma solução para monitorar, controlar e otimizar o uso da energia elétrica. A capacidade do sistema NILM de identificar o consumo individual sem a necessidade de instalar sensores específicos em cada aparelho permite uma visão detalhada do perfil de consumo da casa. Com essas informações, o SGE pode tomar decisões inteligentes sobre como e quando utilizar certos dispositivos para maximizar a eficiência energética e oferecer conforto e conveniência aos usuários. A implementação eficaz exige integração entre o NILM e os dispositivos inteligentes da residência. Nesta e em outras aplicações, o artigo destaca a necessidade de privacidade e segurança dos dados dos usuários, visto que o sistema terá acesso a informações detalhadas sobre os hábitos de consumo da casa.

- **Gestão de Demanda:** Esta aplicação envolve o ajuste do consumo de energia em resposta às condições da rede elétrica, visando otimizar o uso de recursos energéticos e minimizar custos. Por meio da identificação precisa do consumo de energia de diferentes aparelhos em tempo real, a técnica NILM permite que consumidores e gestores de energia implementem estratégias de resposta à demanda mais eficazes. Isso inclui o uso de técnicas conhecidas por desligamento seletivo de cargas, que envolvem o deslocamento do uso de energia para horários de menor demanda, a redução voluntária do consumo durante picos de carga e a participação em programas de resposta à demanda oferecidos pelas concessionárias. Essas ações podem contribuir para a estabilidade da rede e a redução de custos com energia.

- **Deteção de Anomalias:** No contexto dos sistemas NILM, a deteção de anomalias envolve a identificação de padrões de consumo de energia que sinalizam desde fugas até falhas em equipamentos, contribuindo para a manutenção preventiva, eficiência energética e segurança. Anomalias como sobrecargas e curtos-circuitos podem ser identificadas precocemente, reduzindo riscos de danos e acidentes elétricos. Embora promissora, essa aplicação ainda demanda avanços para sua plena utilização, conforme evidenciado por (RASHID, SINGH, *et al.*, 2019).

- **Ambientes de Vivência Assistida:** A técnica NILM pode monitorar e oferecer suporte a independência e segurança de pessoas, especialmente idosos ou pessoas com necessidades especiais. Essa tecnologia permite a identificação de padrões de atividade ou inatividade, sugerindo alterações na rotina diária que podem indicar situações emergenciais ou de risco. Importante destacar que, neste contexto, a

precisão na estimativa de consumo das cargas é secundária em relação à habilidade do sistema em discernir o estado operacional dos aparelhos monitorados (BUCCI, CIANCETTA, *et al.*, 2021). O estudo conduzido por (NOURY, BERENGUER, *et al.*, 2011), através de experimentos de duração de alguns meses, evidenciou não apenas a viabilidade do uso de técnicas NILM em ambientes de vivência assistida, mas também revelou duas descobertas notáveis: primeiramente, a atividade elétrica atua como uma assinatura única da atividade individual, e, secundariamente, tal perfil oferece informações inesperadas sobre o estado de saúde dos moradores. Este trabalho reconhece o vasto potencial dessa aplicação, bem como as amplas possibilidades para futuras investigações na área.

2.1.5.2 Aplicações para Fornecedores de Produtos e Serviços

Além das aplicações voltadas aos consumidores, o autor (HART, 1992) identifica as concessionárias de energia e outros fornecedores de produtos e serviços como beneficiários diretos do NILM, devido à sua capacidade de realizar monitoramento detalhado do consumo de energia em larga escala. As aplicações destacadas incluem:

- **Previsão de Demanda e Tarifação:** Permite às concessionárias aprimorarem suas previsões de demanda com maior precisão, ajustando a produção energética às necessidades reais, otimizando a gestão de recursos e reduzindo custos operacionais. A análise do consumo pode fundamentar modelos de tarifação que reflitam os padrões de consumo individuais.
- **Auditoria Energética:** A auditoria energética é um processo essencial para compreender e otimizar o consumo de energia em edificações, envolvendo a inspeção, análise e avaliação detalhada dos fluxos de energia. Tradicionalmente, este processo requer a instalação temporária de dispositivos de monitoramento por especialistas para coletar dados sobre o uso de energia, visando identificar oportunidades para a economia energética, seja através do upgrade de equipamentos ou da alteração nos padrões de consumo. A integração da técnica NILM revoluciona esta prática ao proporcionar uma coleta de dados mais eficiente e menos intrusiva. Como destacado por (BERGES, GOLDMAN, *et al.*, 2010), o NILM oferece informações granulares sobre o consumo dos equipamentos sem necessidade de instalação de múltiplos sensores, facilitando a identificação de áreas para melhorias

energéticas. Além disso, a capacidade do NILM de fornecer *feedback* em tempo real sobre o impacto das mudanças implementadas permite um processo de otimização contínua, tornando as auditorias energéticas mais dinâmicas e efetivas.

- **Desenvolvimento de Políticas Públicas e Design de Aparelhos:** As informações de consumo energético auxiliam na criação de políticas de eficiência energética e sustentabilidade. Fabricantes de aparelhos podem utilizar esses dados para desenvolver produtos que melhor atendam às necessidades dos consumidores e aos critérios de sustentabilidade.

- **Vantagens Operacionais:** O autor (HART, 1992) também destaca as vantagens da simples instalação e custo reduzido, viabilizando o monitoramento em larga escala sem intrusão nas instalações dos consumidores das concessionárias ou necessidade de permissão explícita, o que facilita a logística de implementação.

-

2.2 CARACTERIZAÇÃO DO PROBLEMA

A potência agregada em uma residência é influenciada diretamente pelo estado operacional das suas cargas em um determinado momento. Para descrever esse processo, define-se a variável $a_i(t)$, um vetor booleano de tamanho n , que indica o estado de cada uma das n cargas em um instante t :

$$a_i(t) = \begin{cases} 1, & \text{se o aparelho estiver ligado} \\ 0, & \text{se o aparelho estiver desligado} \end{cases}$$

para $i = 1, \dots, n$.

A potência total consumida numa residência em qualquer instante t pode ser então expressa como:

$$P_{Total}(t) = \sum_{i=1}^n a_i(t)P_i + e(t) \quad 1$$

onde $P_{Total}(t)$ representa a potência agregada, P_i é a potência consumida pelo aparelho i , e $e(t)$ as contribuições de potência de aparelhos não considerados e o ruído na medição (HART, 1992).

A técnica NILM explorada neste trabalho tem como objetivo solucionar a Equação 1. Este processo envolve a determinação dos estados operacionais das cargas que, ao serem combinados, correspondem à potência agregada medida, simultaneamente buscando minimizar qualquer erro associado a esta reconstrução. Este desafio é formulado como um problema de CO, conforme proposto por (HART, 1992):

$$\hat{a}(t) = \arg_a \min \left| P_{Total}(t) - \sum_{i=1}^n a_i(t)P_i \right| \quad 2$$

onde $\hat{a}(t)$ é um vetor booleano representando o conjunto de estados a_i das cargas que resultam no mínimo erro ou discrepância entre o consumo agregado medido no instante t , $P_{Total}(t)$, e o somatório dos consumos estimados para todos os aparelhos individuais $\sum_{i=1}^n a_i(t)P_i$.

Várias dificuldades são reconhecidas na busca por uma solução direta da Equação 2, tornando-a impraticável a menos que o número de cargas envolvidas seja significativamente pequeno. Uma questão central é a incompletude do conhecimento sobre o conjunto de potências P_i . Em um ambiente residencial, a chegada de novos dispositivos e a retirada de outros ocorre por diversos motivos, e a aplicação da Equação 2 em situações com cargas desconhecidas levaria à tentativa de explicar seu comportamento pela combinação das cargas já conhecidas. Adicionalmente, mesmo na hipótese de todas as cargas serem conhecidas, pequenas variações na potência agregada $P_{Total}(t)$ podem ser interpretadas erroneamente como alterações significativas no estado dos aparelhos. Assim, a solução mais eficaz para o problema de otimização combinatória reside no emprego de um conjunto de heurísticas que são capazes de oferecer soluções adequadas e razoavelmente precisas na maioria dos casos (HART, 1992).

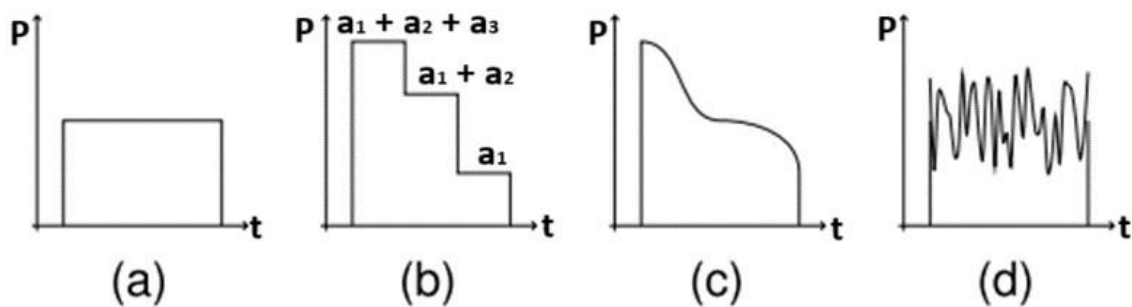
Neste sentido, (HART, 1992) propõe o princípio fundamental da técnica NILM, denominado *switch continuity principle* (SCP — Princípio da Unicidade de Eventos), estabelecendo que em um breve intervalo de tempo espera-se que apenas um pequeno número de cargas (idealmente uma) altere seu estado. Hart validou este princípio através de experimentos que demonstraram que eventos simultâneos ou quase simultâneos ocorreram em apenas 4% das ocasiões, em um intervalo de 2 a 3

segundos. Essa observação, apesar de variar de acordo com cada situação, reduz o risco de interpretar incorretamente pequenas flutuações como grandes mudanças nos estados dos aparelhos. Além do SCP, outras técnicas complementares, apresentadas na Seção 2.7.3 são empregadas para reforçar a eficácia do processo de desagregação.

2.3 TIPOS DE CARGAS

Antes de descrever as etapas de um sistema NILM em detalhes é importante caracterizar e agrupar as cargas elétricas de maneira conceitual. Em seu trabalho, (HART, 1992) propõe quatro modelos que usualmente servem de base para definição das assinaturas elétricas, formalmente apresentadas na Seção 2.4. A Figura 3 ilustra as formas de onda características de cada modelo.

Figura 3 — Ilustração dos modelos de tipos de cargas elétricas proposta por (HART, 1992). (a) Tipo Liga/Desliga, (b) Máquina de Estados Finitos, (c) e (d) Continuamente Variável.



FONTE: Adaptado de (SCHIRMER, MPORAS, 2023).

2.3.1 Tipo 1: Cargas Liga/Desliga

As cargas do tipo Liga/Desliga representam a maior parte das cargas presentes nas residências e se caracterizam pela simplicidade de acionamento e pelos dois estados de funcionamento: ligado e desligado. Sua demanda é aproximadamente constante durante o uso, o que simplifica a desagregação destas cargas. A Figura 3 (a) ilustra esse tipo de carga. Exemplos: torradeiras, aquecedores, lâmpadas e ar-condicionado convencional.

2.3.2 Tipo 2: Cargas com Estados Finitos

As cargas do tipo Estado Finito, ilustradas na Figura 3 (b), possuem estados de operação que se alternam em um ciclo periódico bem definido. Sua desagregação é mais complexa do que a das cargas do tipo Liga/Desliga e apresenta desafios adicionais para sistemas de monitoramento, pois o sistema precisa identificar não apenas a presença da carga na rede, mas também o estado em que ela se encontra. Esse tipo de carga é uma generalização que engloba as cargas do tipo 1. Exemplos: máquinas de lavar e secadoras de roupas, geladeiras com degelo automático e liquidificadores.

2.3.3 Tipo 3: Cargas com Variação Contínua

As cargas do tipo Variação Contínua, ilustradas na Figura 2 (c) e (d), apresentam consumo variável ao longo do tempo, sem padrões cíclicos. Elas representam o maior desafio para a desagregação, pois não possuem características que facilitem a identificação. Exemplos: *dimmers*, ferramentas, TVs e caixas de som.

2.3.4 Tipo 4: Cargas Permanentes

As cargas permanentes, como alarmes e fontes externas, estão sempre presentes na rede elétrica. Elas podem ser consideradas como uma constante no consumo de energia e representam aproximadamente 15 % do consumo total de uma residência (CARRIE ARMEL, GUPTA, *et al.*, 2013). Também são consideradas um desafio para a desagregação, pois não geram eventos característicos ao longo do tempo.

2.4 ASSINATURAS ELÉTRICAS

No cerne da técnica NILM estão as assinaturas elétricas: características mensuráveis que distinguem o consumo das cargas presentes na rede. As assinaturas elétricas fornecem uma "impressão digital" única para cada carga, permitindo sua identificação ao serem capturadas por um sistema NILM. Este capítulo discute as

principais classes de assinaturas, conforme proposto por (HART, 1992): regime permanente e transitórios, explorando suas características, vantagens e desvantagens e aplicabilidade na identificação e monitoramento eficiente das cargas.

2.4.1 Regime Permanente

As assinaturas elétricas de regime permanente são características mensuráveis do consumo de energia que apresentam estabilidade ao longo do tempo. Essas assinaturas compreendem padrões estáveis em grandezas como potência ativa, corrente e admitância. Segundo os autores (FAUSTINE, MVUNGI, *et al.*, 2017) e (CARRIE ARMEL, GUPTA, *et al.*, 2013) a maioria das aplicações NILM na literatura utiliza características de regime permanente para identificação de cargas. A principal vantagem encontra-se na simplicidade do equipamento de aquisição e processamento dos sinais necessários para deste tipo de assinatura, que não requer elevada taxa de amostragem. Nesse sentido, a potência ativa é a métrica de maior relevância para a área devido à sua importância comercial: em residências essa métrica está diretamente relacionada ao faturamento do consumo e em muitos lugares os equipamentos de aquisição necessários já se encontram instalados.

O autor (HART, 1992) identifica três razões principais para a preferência por assinaturas em regime permanente em seus protótipos:

1. Facilidade de detecção: Padrões de consumo estáveis são mais simples de coletar e processar;

2. Universalidade de Ocorrência: Todas as mudanças de estado geram assinaturas em regime permanente. Em contraste, assinaturas transitórias são geradas somente no momento de ativação das cargas, sem um perfil distinto ao desligá-las;

3. Propriedade Aditiva: Embora a identificação de múltiplas cargas acionadas simultaneamente possa ser desafiadora, as assinaturas em regime permanente mantêm a propriedade aditiva, permitindo, com esforço analítico apropriado, a sua distinção. (Detalhes adicionais na Seção 2.6.1).

Embora as assinaturas de regime permanente sejam mais fáceis de detectar, muitas cargas podem apresentar assinaturas semelhantes em uma mesma grandeza, dificultando a identificação individual. Nesses casos, por exemplo, se duas cargas possuem eventos de potências ativas parecidas, recomenda-se ampliar o conjunto de

métricas observadas pelo sistema NILM. O objetivo é estabelecer quais propriedades em seu consumo são capazes de diferenciá-las. É importante notar que este esforço para melhorar o desempenho da desagregação também aumenta o esforço computacional da aplicação, o que pode aumentar seu custo.

Para a desagregação das cargas no sistema NILM desenvolvido neste trabalho, são utilizadas apenas as variações (eventos) de potência ativa como assinatura. Sugere-se o uso da potência reativa como segunda métrica observada quando há a presença de cargas com padrões de consumo de potência ativa similares (CARRIE ARMEL, GUPTA, *et al.*, 2013).

Uma importante consideração deve ser feita neste momento. Numa rede elétrica residencial, a tensão fornecida pode variar $\pm 5\%$ do seu valor nominal, conforme o (“Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST”) . Rápidas flutuações da tensão de fornecimento são normais e podem ser causadas pelo acionamento de outra carga na mesma rede ou a comutação de transformadores. Uma carga linear ligada à rede elétrica cuja tensão possa sofrer essas variações também terá um consumo de corrente com variação na mesma proporção. Desta forma, a potência consumida por essa carga pode sofrer uma variação de quase 10% causada por fatores externos. Uma assinatura elétrica com essa característica não é ideal para um sistema NILM (HART, 1992). Para minimizar a influência das variações de tensão na rede elétrica residencial, (HART, 1992) propõe a utilização de um artifício para a correção da potência ativa com referência à tensão nominal de fornecimento, da seguinte forma

$$P_{Norm}(t) = V_{Nom}^2 \cdot Y(t) = \left(\frac{V_{Nom}}{V(t)} \right)^2 \cdot P(t) \quad 3$$

Onde $P_{Norm}(t)$ é a potência corrigida, V_{Nom} é a tensão nominal de fornecimento da rede, $V(t)$ é a tensão medida ao longo do tempo e $Y(t) = \frac{P(t)}{V^2(t)}$ é a admitância total observada a partir do ponto medido. Quando possível, a potência ativa normalizada $P_{Norm}(t)$ pode ser usada pelo sistema NILM para a identificação das cargas, compensando variações na rede de fornecimento.

2.4.2 Transitórios

Os transitórios são rápidas oscilações que ocorrem nas variáveis elétricas antes de se atingir um valor estável de consumo de energia (regime permanente). O uso de assinaturas transitórias para a identificação de cargas está associado ao uso de equipamentos mais precisos e caros pois são mais difíceis de detectar. Sua utilização se justifica em situações cujas cargas de interesse possuem assinaturas de regime permanente idênticas (HART, 1992). Os transitórios são altamente dependentes do tipo de carga. Dados como a trajetória da curva de corrente, seu tamanho, duração e harmônicos presentes no período de transição também compõe este tipo de assinatura (ZOHA, GLUHAK, *et al.*, 2012).

2.5 AQUISIÇÃO DOS SINAIS

A etapa inicial em um sistema NILM consiste na aquisição dos sinais elétricos, um processo que envolve a coleta de dados de tensão e corrente por meio de um sistema de conversão analógico-digital. Essa transformação dos sinais para um formato digital permite o cálculo de outras grandezas elétricas, como a potência ativa, fundamentais para a subsequente extração de assinaturas elétricas e desagregação de cargas. A seguir são discutidos alguns aspectos relevantes a essa etapa.

2.5.1 Frequência de Aquisição

A determinação da frequência de aquisição é crucial para o projeto de um sistema NILM, influenciando diretamente tanto a capacidade de desagregação das cargas quanto os requisitos de hardware. Uma frequência de aquisição mais elevada possibilita a captura de detalhes finos das assinaturas elétricas, melhorando a precisão na identificação das cargas, mas implica em um maior volume de dados a ser processado. Isso exige equipamentos de maior capacidade e, por consequência, pode elevar os custos de implementação do sistema.

Conforme indicado por (CARRIE ARMEL, GUPTA, *et al.*, 2013), diferentes frequências de aquisição podem ser aplicadas para identificar variados tipos de cargas. Dados coletados em intervalos horários são adequados para identificar cargas influenciadas por fatores ambientais ou de uso programado, tais como motores de

piscina e iluminação externa, enquanto dados de alta resolução, na faixa de MHz, são capazes de distinguir até 100 cargas distintas, incluindo aquelas com assinaturas harmônicas específicas inacessíveis em frequências mais baixas. Neste contexto, cargas do tipo 4, permanentes, exemplificam a necessidade de alta frequência de aquisição, uma vez que apenas suas harmônicas em regime permanente são capazes de identificá-las.

O autor (HART, 1992) observou que a coleta de dados em frequências inferiores a um segundo pode levar à combinação de eventos distintos em uma única amostra, complicando e aumentando o trabalho do algoritmo de desagregação (detalhes adicionais na Seção 2.6.1). No entanto, assinala que aumentar a frequência de aquisição pode minimizar esse problema somente se as cargas apresentarem transitórios breves, pois eventos transitórios prolongados e quase simultâneos tendem a se sobrepor no sinal agregado, independentemente da taxa de amostragem.

A frequência dos dados pode variar entre os sinais adquiridos e as métricas internamente calculadas, e técnicas como o *downsampling* são frequentemente empregadas para reduzir o volume de dados armazenados e os requisitos de processamento, conforme explicado por (HAQ, JACOBSEN, 2018). No mercado, existem diversas soluções de medidores residenciais com frequências de aquisição que variam de 60 Hz a 2 kHz, enquanto sistemas NILM de alta frequência, de 10 MHz a 100 MHz, costumam ser customizados e implicam custos elevados devido à complexidade do *hardware* necessário (ZOHA, GLUHAK, *et al.*, 2012).

A escolha da frequência de aquisição adequada depende das cargas específicas a serem monitoradas e dos objetivos do sistema NILM, não existindo um consenso na literatura sobre a frequência ideal. A decisão deve considerar o equilíbrio entre precisão de identificação, capacidade de processamento e custo, adaptando-se às necessidades da aplicação conforme destacado por (HOSSEINI, AGBOSSOU, *et al.*, 2017).

2.5.2 Resolução e Precisão dos Dados

A resolução dos dados adquiridos define a menor variação de consumo energético que pode ser detectada pelo sistema NILM, impactando significativamente o desempenho da desagregação. Uma maior resolução permite a identificação precisa de variações sutis no consumo de energia, facilitando a distinção entre cargas

pequenas ou com perfis de consumo similares. Geralmente, sistemas que operam com frequências de aquisição superiores a 1 Hz buscam resoluções mínimas de 0,1 W para desagregar eficientemente cargas de menor porte, como pequenos dispositivos eletrônicos. No entanto, muitos sistemas funcionam bem com resoluções mais baixas, da ordem de 10 W, sendo que a desvantagem está na incapacidade de desagregar pequenas cargas (CARRIE ARMEL, GUPTA, *et al.*, 2013).

A presença de ruídos nos dados medidos é outro aspecto que deve ser considerado. A adoção de frequências de aquisição e resoluções mais elevadas pode aumentar a suscetibilidade do sistema ao ruído, tanto ambiental quanto de amostragem, complicando a análise dos dados.

A precisão, por sua vez, refere-se à conformidade dos valores medidos com os valores reais de consumo. Um estudo (“Evaluation of Advanced Metering System (AMS) Deployment in Texas Navigant Consulting (PI) LLC”, 2010) revelou que a precisão média dos medidores comercialmente disponíveis é de 99,96 % com uma margem de erro de ± 2 %. Para a maioria das aplicações, incluindo o faturamento de consumo energético, uma precisão de até 99,5 % é considerada satisfatória. A precisão é um atributo crítico para medidores em sistemas NILM, influenciando diretamente a confiabilidade dos dados coletados e, por extensão, a precisão da desagregação de cargas.

Portanto, a seleção da frequência de aquisição, resolução e a garantia de precisão adequadas são fundamentais para otimizar o desempenho do sistema NILM, exigindo uma análise detalhada das necessidades específicas da aplicação e ponderação entre os benefícios da alta resolução e os desafios técnicos e financeiros associados.

2.6 EXTRAÇÃO DE PROPRIEDADES

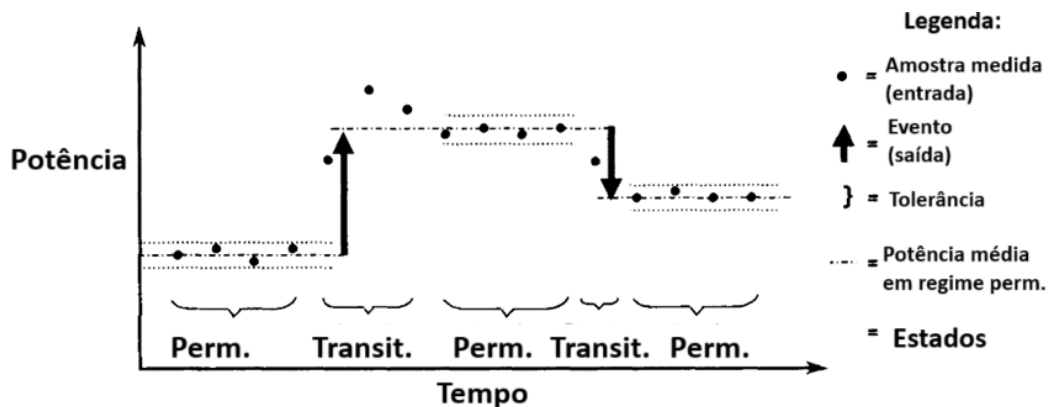
Nesta etapa do sistema NILM são utilizados algoritmos de processamento dos sinais coletados na etapa anterior com o objetivo de identificar mudanças no consumo de energia que caracterizem as cargas. Existem dois métodos fundamentais propostos por (HART, 1992) para essa finalidade, denominados eventos e estados. Eventos são caracterizados como instantes de mudança entre os estados operacionais das cargas, marcando transições significativas no consumo, enquanto os estados representam períodos de consumo estável. Ambos os métodos,

detalhados nas subseções subsequentes, estabelecem a base para sistemas NILM e se mantêm relevantes nos sistemas modernos conforme evidenciado por (FAUSTINE, MVUNGI, *et al.*, 2017).

2.6.1 Métodos Baseados em Eventos

Nesta abordagem, algoritmos de detecção de bordas são empregados para identificar mudanças significativas — eventos de subida e descida — no perfil de consumo energético, correspondentes às transições entre diferentes estados operacionais das cargas. Esses eventos, definidos tanto por sua magnitude quanto pelo momento de ocorrência, são utilizados para a formação das assinaturas energéticas nas etapas subsequentes. Essa abordagem é muito vantajosa pois otimiza o problema da desagregação ao converter um vasto volume de dados numa lista gerenciável de eventos para análise posterior. A Figura 4 ilustra os conceitos de eventos e estados, correlacionando-os com as amostras medidas.

Figura 4 — Ilustração conceitual de eventos de subida e descida e estados transitórios e em regime permanente.



Fonte: Adaptado de (HART, 1992).

A abordagem baseada em eventos permanece relevante e é a mais utilizada em aplicações NILM modernas (BANDEIRA, MARTINS, 2020). A robustez do algoritmo de detecção em relação aos estados transitórios é um requisito fundamental. Transitórios prolongados podem confundir o algoritmo e induzir a identificação equivocada de estados permanentes intermediários, resultando na detecção de múltiplos eventos de menor magnitude, inexistentes, enquanto o sinal transiciona

entre estados reais (HART, 1992). O uso de técnicas como *downsampling*, apresentada na Seção 2.5.1, é uma estratégia eficaz para mitigar essas ocorrências, simplificando a análise dos sinais sem comprometer a precisão na identificação de eventos verdadeiros.

2.6.2 Métodos Baseados em Estados

A abordagem baseada em estados distingue-se pela sua capacidade de modelar cada carga como uma máquina de estados probabilísticos, onde os níveis de consumo energético, tais como o nível de potência ativa, são correlacionados com as probabilidades de transição entre os estados operacionais e o tempo de permanência em cada estado. Este conceito fundamenta-se na observação de que as cargas ao alternarem entre estados operacionais — como por exemplo o ciclo de lavagem e centrifugação de uma máquina de lavar —, manifestam padrões de transição e duração dos estados que são distintos e identificáveis. Esta característica permite a identificação e diferenciação entre cargas com base em seus padrões comportamentais específicos.

Embora robusta, esta metodologia enfrenta desafios específicos decorrentes da necessidade de um conhecimento detalhado sobre a operação das cargas. A construção de um banco de dados abrangente e de longa duração é essencial. Também há a dificuldade de expandir o banco de dados para cargas similares, como aparelhos de diferentes marcas e modelos, pois cada carga possui um conjunto único de estados (FAUSTINE, MVUNGI, *et al.*, 2017).

2.7 ALGORITMOS DE CLASSIFICAÇÃO

A etapa final de um sistema NILM tem como objetivo principal desagregar o sinal elétrico agregado, ou seja, identificar e classificar quais cargas individuais estão ativas na rede elétrica a partir das amostras coletadas no PCC. Esse processo pode ser visto como uma engenharia reversa, onde se comparam os dados agregados com um modelo pré-definido ou aprendido ao longo do tempo, atuando como uma biblioteca contendo as assinaturas elétricas das cargas. Para realizar essa tarefa complexa, algoritmos de aprendizado de máquina são utilizados. Neste contexto, os

algoritmos de aprendizado de máquina aplicados podem ser categorizados em dois tipos principais: supervisionados e não supervisionados, detalhados a seguir.

2.7.1 Supervisionados

Os algoritmos supervisionados exigem dados específicos de cada carga para realizar a desagregação. Essa abordagem opera em duas fases principais: treinamento e desagregação. Durante a fase de treinamento, o objetivo é construir um modelo que contenha as assinaturas elétricas de cada carga, geralmente realizado de forma intrusiva através de sensores individuais. Na fase de desagregação, o algoritmo utiliza os eventos detectados nos sinais agregados coletados no Ponto de Conexão Comum (PCC) para identificar as cargas.

A exigência de um banco de dados detalhado para cada carga apresenta um desafio significativo, tornando esta fase mais complexa e demorada conforme o número de cargas aumenta. A portabilidade do modelo entre diferentes residências é limitada devido à variação das cargas, afetando a escalabilidade desta abordagem (KHAZAEI, STANKOVIC, *et al.*, 2020). Além disso, mudanças no consumo de energia devido ao desgaste das cargas podem comprometer a precisão e a robustez do modelo ao longo do tempo (LIU, CHEN, 2014). No entanto, o avanço na pesquisa desta área e o desenvolvimento de bancos de dados abrangentes têm facilitado a criação de algoritmos NILM supervisionados, oferecendo uma base de dados rica para o treinamento (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022). Detalhes adicionais dos bancos de dados são discutidos na 2.10.

A literatura recente destaca o potencial de algoritmos supervisionados como Redes Neurais Artificiais e *k-Nearest Neighbors*. O trabalho (GOPINATH, KUMAR, *et al.*, 2020) apresenta uma análise abrangente de algoritmos NILM modernos. Já o trabalho (REHMAN, LIE, *et al.*, 2021) compara o desempenho de 10 algoritmos avançados. Este trabalho aborda uma aplicação do algoritmo CO, precursor das abordagens mais avançadas. A Seção 2.7.3 apresenta e detalha este algoritmo.

2.7.2 Não Supervisionados

Os algoritmos não supervisionados dispensam a necessidade de treinamento prévio com assinaturas elétricas de cada carga. Ao invés disso, esses algoritmos

aprendem diretamente a partir dos dados agregados, adaptando-se continuamente ao longo do tempo. Esta abordagem busca resolver alguns desafios encontrados nos métodos supervisionados, como a degradação de cargas, a adição de novos dispositivos ao ambiente e a transferência de soluções NILM entre diferentes residências, melhorando assim a escalabilidade do sistema (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022). Contudo, esse estudo ressalta que o desempenho dos algoritmos não supervisionados tende a ser inferior quando comparado aos resultados obtidos por métodos supervisionados. Destacam-se, entre as abordagens não supervisionadas, implementações que se baseiam nos algoritmos *Hidden Markov Models* e *k-means*.

Em linha com as tendências atuais em tecnologia de algoritmos, pesquisas recentes na área NILM apontam para o desenvolvimento de abordagens híbridas, combinando as vantagens de ambas as classes de algoritmos discutidas anteriormente, dando origem a uma nova categoria conhecida como algoritmos semi-supervisionados. O trabalho de (KASELIMI, PROTOPAPADAKIS, *et al.*, 2022) apresenta essa inovação.

2.7.3 O Algoritmo CO

O algoritmo de desagregação CO, introduzido por (HART, 1992), estabeleceu-se como um paradigma na área NILM, servindo de *benchmark* para a avaliação de abordagens mais avançadas e modernas (SCHIRMER, MPORAS, 2023). Categorizado como um método supervisionado, o algoritmo CO opera em fases distintas de treinamento e desagregação. A Figura 5 destaca um fluxograma que resume o processo deste algoritmo, enfatizando cada fase e sua contribuição para as etapas gerais de um sistema NILM. A descrição detalhada feita por (HART, 1992) é crucial para a compreensão dos mecanismos do algoritmo e a consulta ao artigo original é fortemente recomendada. Além disso, os experimentos conduzidos neste estudo são baseados em uma versão do algoritmo CO proposto por (HART, 1992) implementado no *software* NILMTK, conforme descrito na Seção 2.8. Segue-se uma apresentação do algoritmo abordando concomitantemente alguns detalhes pertinentes à sua implementação neste *software*.

2.7.3.1 Treinamento

A Figura 5 (b) mostra os quatro passos do processo de treinamento, descritos a seguir:

1º. Aquisição dos Sinais Individuais: A base para o treinamento do algoritmo de desagregação é a coleta intrusiva dos sinais elétricos de cada aparelho. Neste estudo, os experimentos são conduzidos utilizando exclusivamente sinais de potência ativa (W) para análise das cargas.

2º. Pré-Processamento: O tratamento inicial dos sinais coletados é realizado para eliminar anomalias e ruídos, empregando uma técnica de *downsampling* que consolida as amostras em blocos de um minuto, calculando a potência média dentro desses intervalos. Este é o menor período das amostras que o *software* NILMTK consegue manipular para realizar a desagregação.

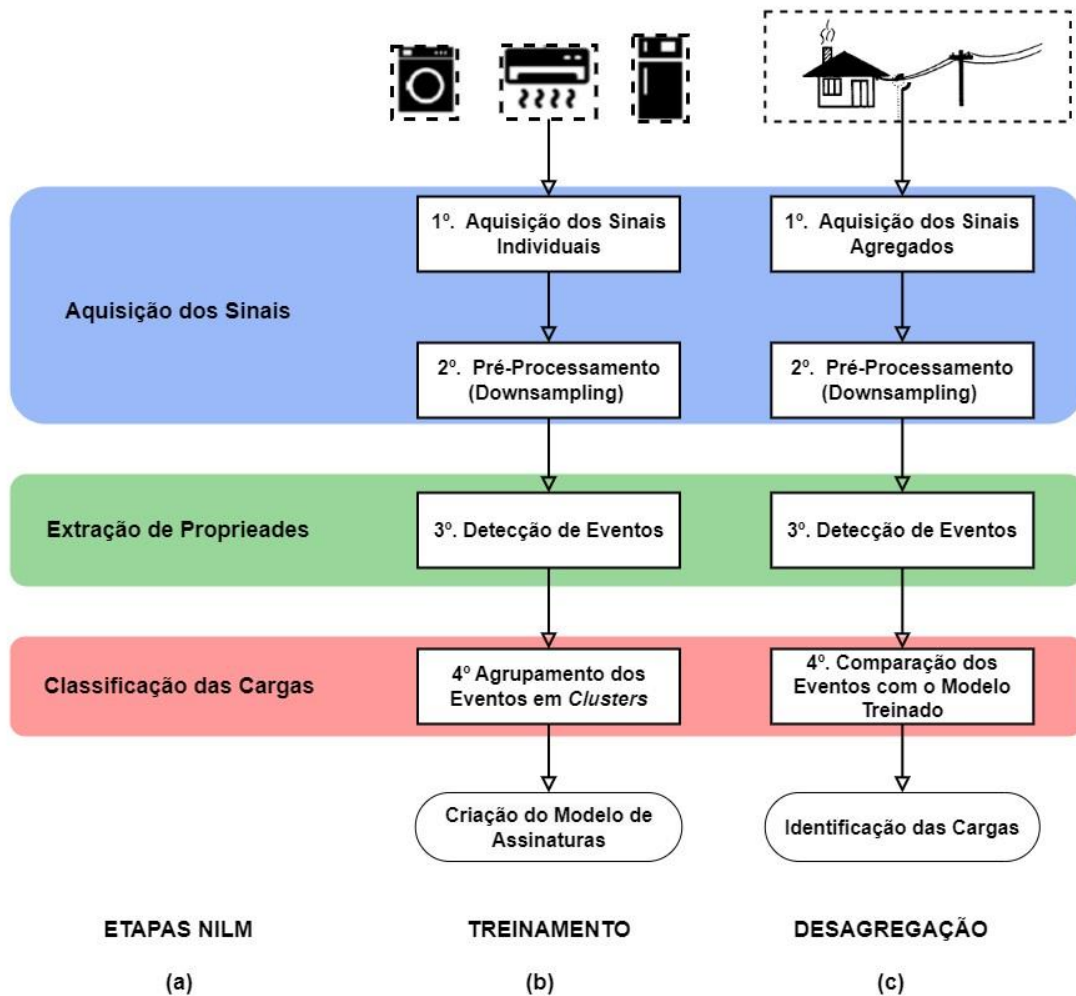
3º. Detecção de Eventos: Essa fase envolve a identificação de variações significativas na série temporal de consumo, as quais podem sinalizar eventos de comutação. Um registro detalhado desses eventos é compilado, incluindo a magnitude e o instante de ocorrência.

4º Agrupamento dos Eventos em *Clusters*: Os eventos detectados para cada carga são organizados em *clusters* em um espaço de assinaturas elétricas com n dimensões, onde n é determinado pelo número de assinaturas observadas. A Figura 6 ilustra esse processo, em que os *clusters* representam regiões de concentração de pontos característicos de diferentes aparelhos, em termos de potência reativa (VAr) e potência ativa (W). É importante ressaltar que, para o propósito específico deste estudo, o espaço de assinaturas é unidimensional, limitando-se à análise da potência ativa dos sinais elétricos.

A técnica de *clusterização* é fundamental na etapa de classificação das cargas, pois permite a identificação precisa de qual dispositivo é responsável por cada evento de consumo de energia. A distribuição dos *clusters*, como ilustrada na Figura 6, revela a diversidade de padrões de consumo elétrico. As variações nas formas e dimensões dos *clusters* refletem as características operacionais distintas dos aparelhos, como, por exemplo, a formação de *clusters* compactos por dispositivos com cargas primariamente resistivas, e de *clusters* mais dispersos por dispositivos com cargas indutivas. Adicionalmente, o gráfico destaca um desafio previamente mencionado: a sobreposição de *clusters* para dispositivos que compartilham assinaturas elétricas

semelhantes, exemplificado pelo agrupamento próximo de diferentes tipos de lâmpadas.

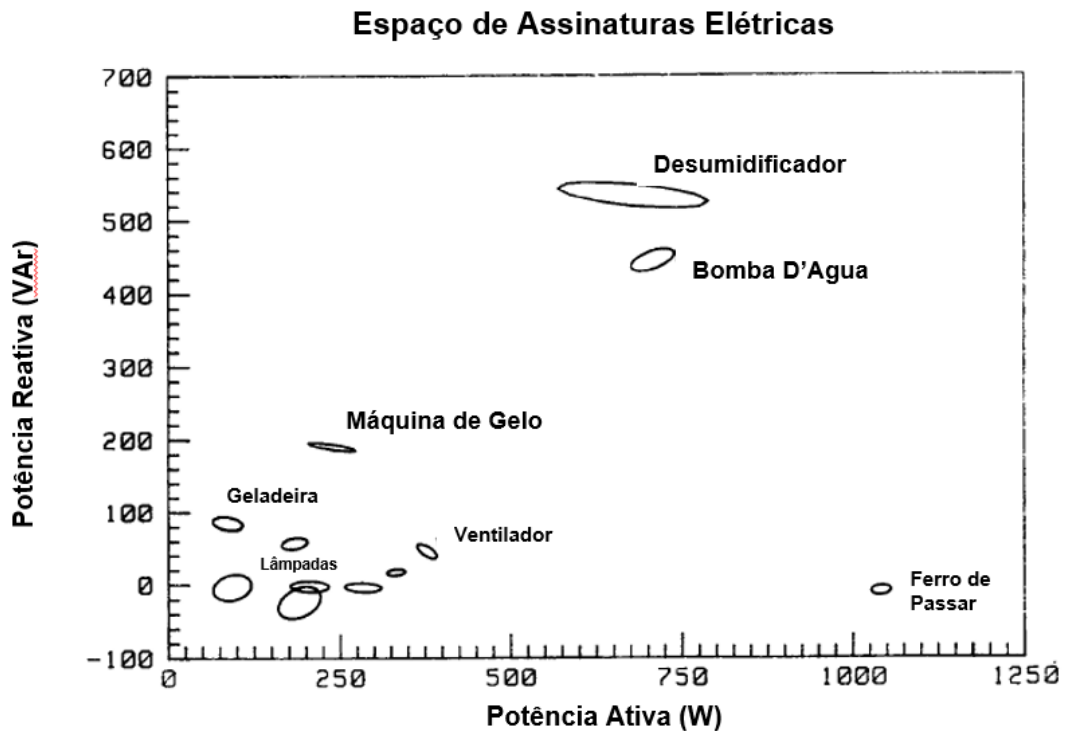
Figura 5 — Fluxograma do Algoritmo CO, correlacionando: as três etapas de um sistema NILM (a), fase de treinamento (b) e fase de desagregação (c).



Fonte: Elaborado pelo Autor (2023), adaptado de (HART, 1992).

Finalizando o treinamento, o algoritmo consolida as informações de todas as cargas analisadas resultando em um modelo de assinaturas elétricas. Este modelo é fundamental para a desagregação subsequente, onde o CO decompõe a potência agregada em componentes individuais. Este modelo, composto pelos *clusters* identificados, serve como base para a subsequente fase de desagregação, onde o algoritmo CO aplica o conhecimento adquirido para decompor a potência agregada em suas componentes individuais.

Figura 6 —Representação do modelo de assinaturas elétricas formado por *clusters* de potências ativa e reativa.



Fonte: Adaptado de (HART, 1992).

2.7.3.2 Desagregação

A Figura 5 (c) mostra os cinco passos do processo de desagregação. Os três primeiros passos: **Aquisição dos Sinais Agregados**, **Pré-Processamento** e **Deteção de Eventos** são essencialmente iguais ao processo realizado durante o treinamento, com a diferença de que neste momento, o sistema captura e trata os dados de potência agregada, coletados no PCC, onde o consumo de todas as cargas se sobrepõe.

No 4º passo, **Comparação dos Eventos com o Modelo Treinado**, os eventos identificados são comparados com o modelo de assinaturas elétricas elaborado durante a fase de treinamento. Essa comparação é feita ao representar os eventos como vetores de potência, que são então projetados no espaço de estados definido pelo modelo. A distância entre o vetor de evento e os *clusters* do modelo é calculada, determinando-se a carga cujo *cluster* está mais próximo. Esse processo resulta na atribuição do evento à carga correspondente mais provável.

É importante ressaltar que o algoritmo CO aborda os eventos detectados em cada amostra de maneira independente, o que simplifica significativamente o problema de desagregação, reduzindo sua complexidade computacional. Contudo, essa abordagem pode limitar a capacidade do algoritmo de capturar padrões de consumo a curto e longo prazo ou correlações entre dispositivos distintos. Para aprimorar a assertividade do processo, são aplicadas heurísticas específicas que funcionam como filtros para regular a detecção de eventos próximos no tempo. Uma dessas heurísticas estabelece que, após um evento de ativação ser detectado para uma carga, deve-se esperar um minuto antes de considerar a possibilidade de um novo evento para a mesma carga. De forma similar, um intervalo de dois minutos é necessário após um evento de desativação antes que um evento de ativação subsequente possa ser reconhecido para a carga em questão. Estes intervalos são uma forma de aumentar a robustez do algoritmo e reduzir a probabilidade de erros.

Outra abordagem já citada na Seção 2.2 é a adoção do SCP, que pressupõe que cada evento é causado por uma única mudança de estado em um aparelho. No entanto, eventos sobrepostos podem ocorrer, apresentando dois cenários desafiadores:

1ª Quando o evento identificado não se alinha exatamente com nenhum *cluster* conhecido, sugerindo a possibilidade de ser a combinação de eventos de múltiplas cargas. Nessa situação, o algoritmo tenta decompor o vetor em componentes que correspondam a eventos de cargas diferentes, enfrentando o desafio de atribuir corretamente os eventos;

2ª Quando o evento está suficientemente próximo a um *cluster* de outra carga, levando a uma atribuição equivocada do evento a essa carga.

Conforme discutido anteriormente na Seção 2.1.3, este é um desafio recorrente em aplicações NILM, que pode ser abordado ao ampliar a dimensão do espaço de assinaturas observadas, melhorando assim a capacidade de distinção entre eventos e cargas.

Por fim, a resolução do problema CO, conforme proposto na Equação 2 e utilizando as técnicas mencionadas, cumpre o objetivo principal da técnica NILM: identificar as cargas individuais que compõem o sinal de consumo agregado.

2.8 SOFTWARE NILMTK

O NILMTK, introduzido por (BATRA, KELLY, *et al.*, 2014), é um *software* de código aberto baseado em Python, criado para incentivar a pesquisa na área NILM. A arquitetura modular do NILMTK oferece um ecossistema completo de ferramentas que abrangem todas as etapas de um sistema NILM e permite a integração de diferentes algoritmos de desagregação, capacitando a investigação de desempenho e comparação de resultados. As principais funcionalidades do *software* incluem:

- **Importação de Dados:** O NILMTK estabeleceu um formato de dados, o NILMTK-DF, especificamente para bancos de dados desenvolvidos nesta área. Esse formato padroniza a estruturação dos dados, permitindo testar os algoritmos em diferentes cenários de cargas.

- **Pré-processamento:** Inclui ferramentas para limpeza, normalização e tratamento dos dados, preparando-os para análise efetiva.

- **Treinamento de Modelos:** Permite o treinamento de diversos algoritmos supervisionados pré-implementados, possibilitando a exploração de diferentes técnicas.

- **Execução de Algoritmos:** Fornece uma interface unificada para a aplicação de diversos algoritmos de desagregação, permitindo testes e comparações diretas entre o desempenho de cada técnica.

- **Avaliação de Desempenho:** Dispõe de um conjunto diversificado de métricas para avaliar a precisão e eficácia dos algoritmos em cada cenário, possibilitando a comparação e seleção de modelos com base em critérios objetivos e reproduzíveis

A ampla adoção do NILMTK pela comunidade de pesquisa NILM decorre da capacidade do *software* em padronizar experimentos e gerenciar grandes volumes de dados de maneira eficiente. Este estudo utiliza a versão 0.4.3 do NILMTK e se concentra especialmente em dois módulos: "*CombinatorialOptimization*", que compreende uma implementação do algoritmo CO proposto por (HART, 1992), e o módulo "*Metrics*", que permite a avaliação do seu desempenho. A próxima seção apresenta em detalhes as métricas utilizadas para mensurar os resultados dos experimentos realizados com o algoritmo CO.

2.9 MÉTRICAS DE DESEMPENHO

A seleção de métricas de desempenho robustas e confiáveis é fundamental para a avaliação precisa e imparcial de algoritmos NILM. As métricas fornecem uma base quantitativa para comparar diferentes técnicas, identificar seus pontos fortes e fracos e, por fim, selecionar a solução mais adequada para cada aplicação específica. Nesse sentido, o melhor algoritmo é aquele que apresenta a melhor capacidade de identificar as cargas do conjunto agregado de sinais, uma vez que este é o objetivo do sistema.

A revisão feita por (PEREIRA, NUNES, 2018) evidencia a falta de consenso na comunidade acadêmica sobre quais métricas devem ser utilizadas para medir e reportar o desempenho de aplicações NILM. Apesar disso, diversas métricas se consolidaram como referências na área e as principais métricas utilizadas para mensurar algoritmos baseados em eventos, de acordo com (ANGELIS, TIMPLALEXIS, *et al.*, 2022), são apresentadas a seguir. Em todos os casos, quanto mais próximo de 1,0 o algoritmo desempenha na métrica avaliada, melhor é seu resultado quanto ao cenário e quanto às variáveis de sucesso de classificação ponderadas.

2.9.1 Precisão

Esta métrica, mostrada na Equação 4, indica o sucesso na classificação de uma determinada classe de eventos durante a ocorrência de todos os eventos verificados.

$$Precisão = \frac{TP}{TP + FP} \quad 4$$

Onde, TP indica o número de vezes que uma classe de eventos foi corretamente identificada (sucesso de classificação) e FP indica o número de vezes que esta mesma classe de eventos foi prevista sem que ela tenha ocorrido (erro de classificação).

2.9.2 Recall

A métrica *Recall*, mostrada na Equação 5, indica o sucesso obtido na classificação de um determinado grupo de eventos em relação ao número de vezes que os eventos dessa classe verdadeiramente ocorreram. O valor unitário dessa métrica mostra que o algoritmo tem habilidade excepcional para reconhecer essa classe de eventos sempre que há sua ocorrência.

$$Recall = \frac{TP}{TP + FN} \quad 5$$

Onde, FN indica o número de vezes que uma classe não foi prevista pelo algoritmo (erro de classificação).

2.9.3 F1-Score

A métrica F1-Score, apresentada na Equação 6, fornece a média harmônica entre as métricas Precisão e *Recall*. Por levar em consideração as duas métricas complementares apresentadas acima, o F1-Score é uma das métricas mais relevantes na área.

$$F1\ Score = 2x \frac{Precisão \times Recall}{Precisão + Recall} \quad 6$$

2.10 BASES DE DADOS

A proliferação de bases de dados públicas de alta qualidade impulsionou significativamente os avanços na área NILM nos últimos anos (ANGELIS, TIMPLALEXIS, *et al.*, 2022). Estes bancos de dados oferecem informações úteis tanto para treinamento quanto para teste de algoritmos de desagregação e são formados por conjuntos de dados que capturam o consumo de energia elétrica em níveis agregado e desagregado, dentro de ambientes residenciais ou comerciais. Cada base possui características próprias tais como número de cargas observadas, frequência de amostragem e tipos de grandezas elétricas registradas.

O artigo (BATRA, PARSON, *et al.*, 2014) apresenta uma extensa revisão da literatura onde são listados 26 bancos de dados públicos. Dentre as múltiplas bases disponíveis, o UK-DALE (*United Kingdom Domestic Appliance Level Electricity*) (KELLY, KNOTTENBELT, 2015) destaca-se pelo seu detalhamento na coleta de sinais elétricos provenientes de 5 residências, cujos dados agregados dos sinais de tensão e corrente elétrica disponibilizados foram medidos na frequência de 16 kHz e os dados de potência ativa e potência aparente, na frequência de 1 Hz. Este banco de dados também disponibiliza dados desagregados de potência ativa de cada aparelho, amostrados na frequência de 1/6 Hz.

Outro banco de dados que se destaca é o REDD (*Reference Energy Disaggregation Data Set*) (KOLTER, JOHNSON, 2011), desenvolvido por engenheiros do MIT e elaborado especialmente para estudos na área NILM. Composto por dados coletados ao longo de várias semanas para 6 residências, fornece valores de potência ativa agregada e desagregada das cargas instaladas. A taxa de amostragem é de 16 kHz para os dados do PCC e 3 segundos para os dados individuais. Para a coleta dos dados foram utilizados sensores comercialmente disponíveis, conforme descrito no seu artigo de lançamento (KOLTER, JOHNSON, 2011). São disponibilizadas algumas informações genéricas sobre as cargas, porém dados como marca, modelo e potência nominal são omitidos em nome da generalização das cargas.

2.11 CONCLUSÃO

Ao longo deste capítulo, apresentaram-se os fundamentos teóricos essenciais à compreensão da técnica NILM, abordando desde a conceituação inicial até as aplicações práticas. A revisão teórica foi iniciada com a obra de George W. Hart, evidenciando sua influência no desenvolvimento da área. Em seguida, a técnica NILM foi explorada em suas várias dimensões, incluindo a introdução à metodologia, a contextualização da motivação e do contexto de aplicação, os princípios operacionais, os desafios técnicos, bem como as classificações dos sistemas NILM e suas aplicações significativas.

Este capítulo destacou também a importância das assinaturas elétricas para a identificação eficaz das cargas, a relevância da precisão na aquisição dos sinais e os desafios associados às etapas de extração de propriedades e classificação das

cargas. Além disso, discutiu-se a aplicação de algoritmos de aprendizado de máquina para aprimorar a técnica NILM, evidenciando o potencial da tecnologia e da inovação na pesquisa aplicada.

Conclui-se, portanto, que a técnica NILM representa uma abordagem promissora para o monitoramento e a gestão do consumo de energia elétrica, oferecendo uma alternativa econômica e eficiente em comparação aos métodos tradicionais de ILM. A evolução da técnica, sustentada pela pesquisa acadêmica e pelo desenvolvimento tecnológico, aponta para a possibilidade de superação dos desafios técnicos existentes e para a ampliação de suas aplicações práticas.

O capítulo subsequente estabelece os procedimentos experimentais e as técnicas analíticas empregadas neste estudo, cujo objetivo é investigar e avaliar o desempenho do algoritmo CO na desagregação de cargas elétricas.

3 METODOLOGIA PROPOSTA

Este capítulo apresenta a metodologia adotada para realizar experimentos de simulação envolvendo a técnica de desagregação dos sinais de potência em um ambiente residencial. Para a execução dos testes, utiliza-se o *software* NILMTK versão 0.4.3, com ênfase em seus módulos “*CombinatorialOptimization*” e “*Metrics*”. O ambiente de execução escolhido para o NILMTK é a Interface de Desenvolvimento Integrado (IDE) Jupyter Notebook. A seleção desses recursos fundamenta-se em sua pertinência e importância para área, conforme discutido em seções anteriores. A seguir, detalham-se as técnicas propostas e as justificativas para sua seleção.

3.1 DEFINIÇÃO DO CONJUNTO DE DADOS UTILIZADOS

Nesta subseção apresentam-se os detalhes e as justificativas para a escolha dos dados analisados.

3.1.1 Banco de Dados

Após cuidadosa consideração, optou-se por focar em cargas e residências reais, em detrimento de um banco de dados sintético. Esta escolha foi influenciada pelo reconhecimento de que a utilização de dados simulados pode comprometer a avaliação de algoritmos de desagregação, conforme destacado por (KLEMENJAK, KOVATSCH, *et al.*, 2020). Dados reais apresentam complexidade intrínseca e ruídos, proporcionando um teste mais rigoroso e fiel das capacidades de desagregação dos algoritmos e contribuindo para resultados mais relevantes.

A seleção do banco de dados para esta pesquisa pautou-se na qualidade e no tipo das informações disponíveis, conforme indicado pela revisão de (BATRA, PARSON, *et al.*, 2014), bem como na sua relevância acadêmica, refletida pelo volume de citações conforme reportado por (IQBAL, MALIK, *et al.*, 2021). Entre as opções consideradas, destacaram-se os seguintes bancos de dados: REDD (KOLTER, JOHNSON, 2011), UK-DALE (KELLY, KNOTTENBELT, 2015) e BLUED (*Building-level fully-labeled dataset for electricity disaggregation*) (ANDERSON, OCNEANU, *et al.*, 2012). Análises mais profundas revelam que, embora o segundo e terceiro banco de dados ofereçam taxa de amostragem compatível com o estudo, as residências

monitoradas apresentam grande quantidade e variedade de aparelhos, incluindo cargas permanentes e continuamente variáveis. Essas características são indesejáveis pois, devido ao seu princípio de funcionamento baseado em eventos, já se sabe que cargas continuamente variáveis e permanentes são um desafio para o algoritmo CO por não apresentarem eventos consistentes.

O banco de dados REDD (KOLTER, JOHNSON, 2011), por sua vez, destacou-se em vários quesitos:

- A qualidade e a resolução dos dados, que incluem medições de potência de potência ativa para cada carga com uma taxa de amostragem de 3 segundos;

- A abrangência do estudo, monitorando 6 residências durante aproximadamente um mês e registrando uma ampla variedade de tipos de cargas instaladas;

- A relevância acadêmica, sendo o mais citado em publicações científicas com mais de 1000 referências (IQBAL, MALIK, *et al.*, 2021).

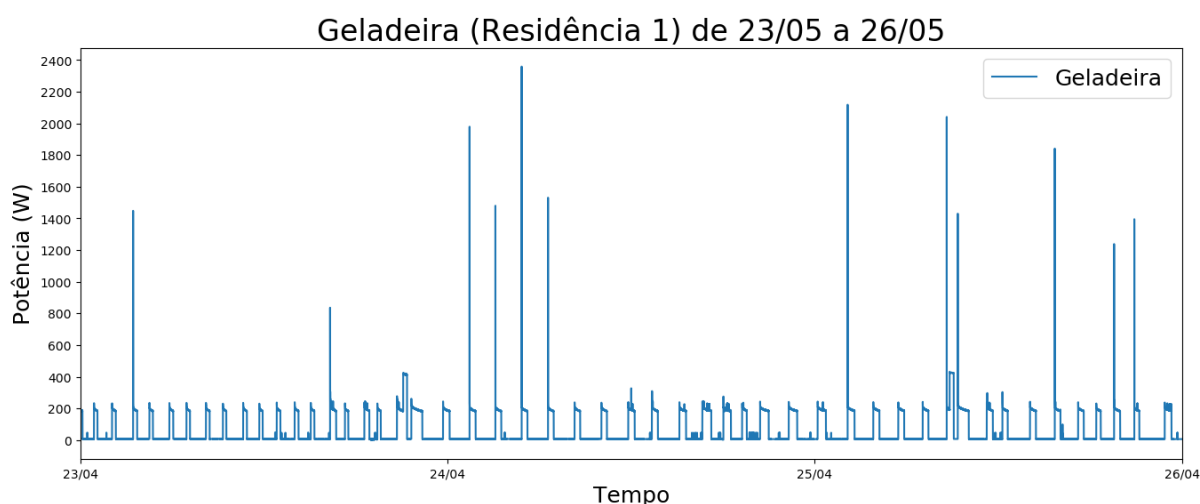
3.1.2 A Residência Selecionada e Suas Cargas Relevantes

A seleção detalhada das seis residências disponíveis no banco de dados REDD indicou que a Residência 1 é a mais adequada para nossa pesquisa, devido à quantidade e a variedade de cargas presentes (16 cargas). Dentre as cargas, algumas têm a mesma finalidade e são diferenciadas pelo número da sua instância. A Tabela 1 resume as cargas instaladas, oferecendo uma visão sobre seus tipos e perfis de consumo, baseando-se em uma análise preliminar de suas curvas de consumo. Vale ressaltar que o banco de dados se limita a fornecer apenas essas informações genéricas, sem detalhar características dos aparelhos como marca, modelo, ano de fabricação ou potência nominal.

A escolha da residência para este estudo foi influenciada significativamente pela presença e pelo tipo de geladeira em cada casa, considerando sua relevância no consumo total em residências brasileiras, como mencionado na Seção 1. A análise focou na Residência 1, que se distinguiu por conter uma geladeira com características de uma carga do tipo Estado Finito, conforme ilustrado nas suas curvas de consumo (Figura 7 e Figura 9), em que é possível notar estados bem definidos. Em contraste, as demais residências ou não possuem dados individuais sobre suas geladeiras ou as unidades monitoradas exibem padrões de consumo de compressores inverter, que

se enquadram no modelo de carga continuamente variável. Tais características desqualificam essas residências da seleção a Residência 1 para o restante deste estudo.

Figura 7 — Consumo de potência ativa da geladeira da Residência 1 ao longo de três dias.

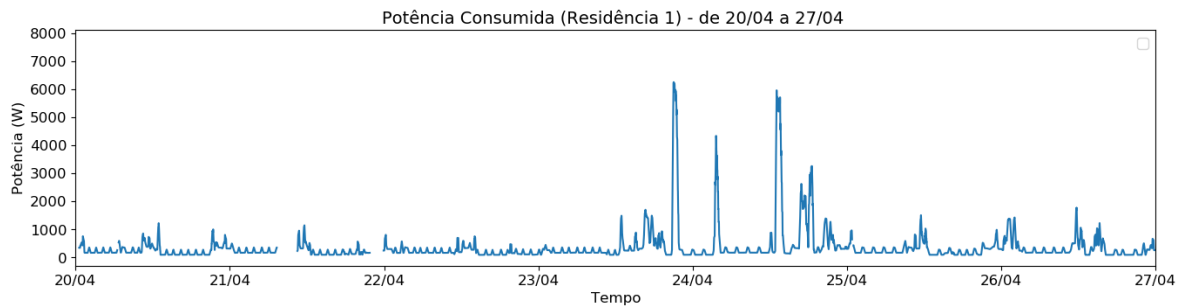


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Os bancos de dados podem apresentar limitações, incluindo falhas e lacunas nas medições, que afetam a integridade dos dados. Essas interrupções podem ser causadas por pausas na coleta, falhas nos dispositivos de medição ou ausência de registros em determinados períodos, comprometendo a representatividade temporal e a precisão das análises. A identificação e o tratamento apropriado dessas lacunas são fundamentais para prevenir interpretações errôneas ou vieses nos resultados.

Nos dados referentes à Residência 1, observam-se lacunas por curtos períodos. No entanto, essas lacunas não impactam significativamente a análise, visto que, durante o intervalo específico de duas semanas, de 18 de abril a 02 de maio de 2011, os dados agregados e individuais mostram consistência, com poucas interrupções. É importante ressaltar que, embora os registros individuais prossigam até 24 de maio, a ausência de dados agregados correspondentes limita a desagregação após este intervalo. A Figura 8 mostra o gráfico da potência agregada medida na Residência 1 durante uma semana.

Figura 8 — Potência agregada medida na Residência 1 entre os dias 20/04 e 27/04/2011.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

As cargas instaladas foram classificadas conforme o seu consumo ao longo do período total disponibilizado (aproximadamente um mês). São apresentadas abaixo, destacando-se as de maior consumo:

1 ^a	Geladeira:	44,7 kWh
2 ^a	Lavadora e Secadora de Roupas:	32,6 kWh
3 ^a	Lâmpadas (1 ^a instância):	30,7 kWh
4 ^a	Tomadas (2 ^a instância):	22,9 kWh
5 ^a	Lava-louças:	19,9 kWh
6 ^a	Lâmpadas (2 ^a instância):	18,2 kWh
7 ^a	Micro-ondas:	16,9 kWh
8 ^a	Tomadas (1 ^a Instância):	16,78 kWh
9 ^a	Lâmpadas (2 ^a Instância):	11,81 kWh
10 ^a	Forno elétrico:	8,81 kWh
11 ^a	Desconhecido (1 ^a Instância):	5,22 kWh
12 ^a	Tomadas (3 ^a Instância):	4,50 kWh
13 ^a	Tomadas (4 ^a Instância):	2,22 kWh
14 ^a	Fogão elétrico:	0,41 kWh
15 ^a	Aquecedor de ambientes:	0,096 kWh
16 ^a	Desconhecido (2 ^a Instância):	0,000085 kWh

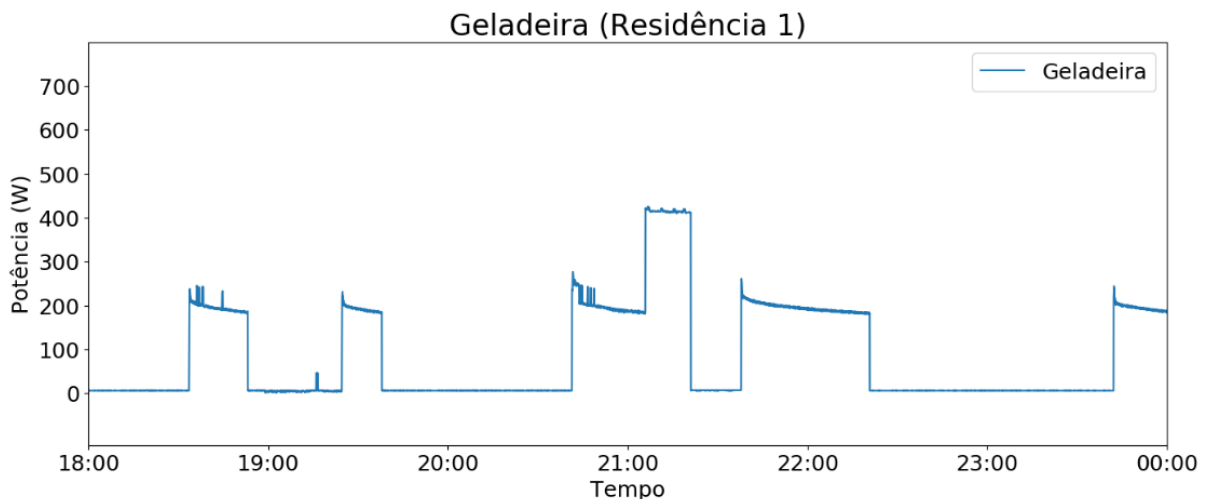
Os valores identificados direcionam a priorização e o exame detalhado dessas cargas nas fases subsequentes. Priorizam-se para desagregação as cargas de maior consumo, alinhando-se ao critério de relevância. Tal estratégia tem o propósito de otimizar a utilidade prática do modelo de desagregação, focando na avaliação dos elementos que mais impactam o consumo total da residência. Adicionalmente, este método permite investigar o efeito de outras cargas não incluídas no conjunto inicial de dados de treinamento.

A seguir são apresentados alguns detalhes das cargas que apresentaram maior consumo de energia.

3.1.2.1 Geladeira

A geladeira se destaca como a principal consumidora de energia na residência. A característica fundamental de seu funcionamento (evidenciada na Figura 8) é a operação intermitente de seu compressor elétrico monofásico. Esse padrão é crucial para otimizar a eficiência energética, permitindo que o aparelho desligue quando a temperatura atinge o nível desejado e ligue novamente conforme necessário. Observa-se, ainda, o ciclo de degelo sistemático que ocorre a cada dois dias. Essa regularidade é uma característica projetada para garantir o bom desempenho do aparelho, removendo o acúmulo de gelo nas bobinas de resfriamento por meio de elementos resistivos.

Figura 9 — Consumo de potência ativa da geladeira da Residência 1, com detalhe para o momento de acionamento do seu sistema de degelo.



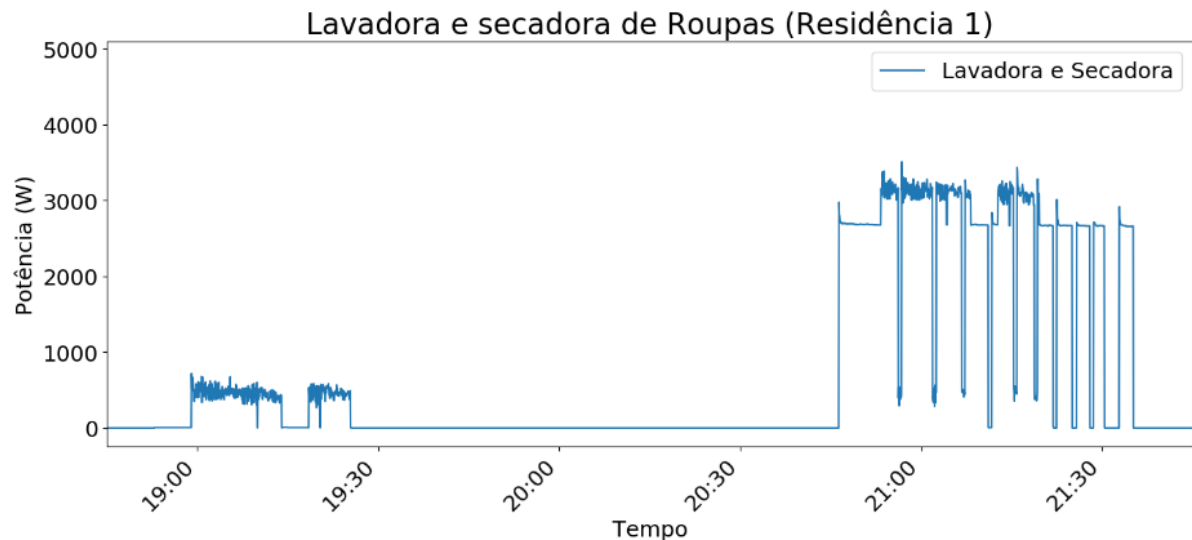
Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

3.1.2.2 Lavadora e Secadora de Roupas

A Figura 10 apresenta o gráfico do ciclo de operação da máquina de lavar e secar roupas na Residência 1, destacando os padrões característicos das diferentes fases do processo. A modelagem dessa carga pode ser feita utilizando o método de

estados finitos, tratando cada fase operacional como um estado distinto. Os eventos relacionados a cada estado são identificados pelo algoritmo nas etapas subsequentes da metodologia. É importante mencionar que o uso da máquina segue uma periodicidade semanal, o que impõe uma restrição à quantidade de dados disponíveis para o treinamento do modelo.

Figura 10 — Consumo de potência ativa ao longo de um ciclo completo da Lavadora e secadora de Roupas.

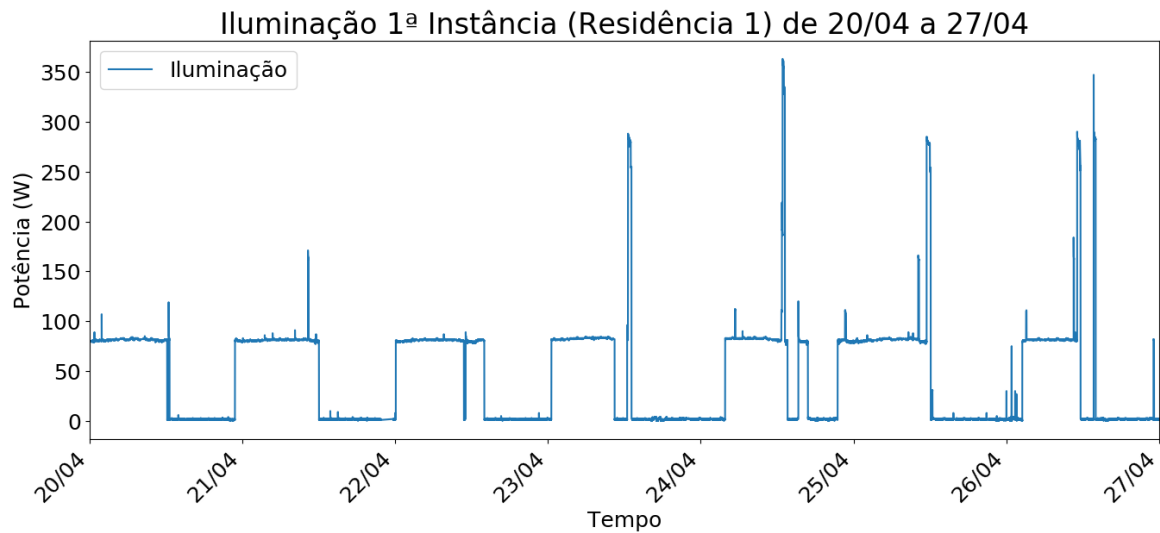


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

3.1.2.3 Circuito de Iluminação

A Figura 11 mostra o consumo de potência ativa do circuito de iluminação (1ª Instância), com comportamentos análogos observados nas demais instâncias, que exibem pequenas variações nos níveis de consumo. É possível modelar esta carga como sendo do tipo Ligado e Desligado. Há outras instâncias que apresentam mais de uma carga no circuito, indicando, por exemplo, se tratar de mais de um cômodo. Nestes casos é possível caracterizá-los como tipo Estado Finito. A Tabela 1 indica em detalhes o tipo de carga presente em cada circuito.

Figura 11 — Consumo de potência ativa do circuito de iluminação (1ª Instância).



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

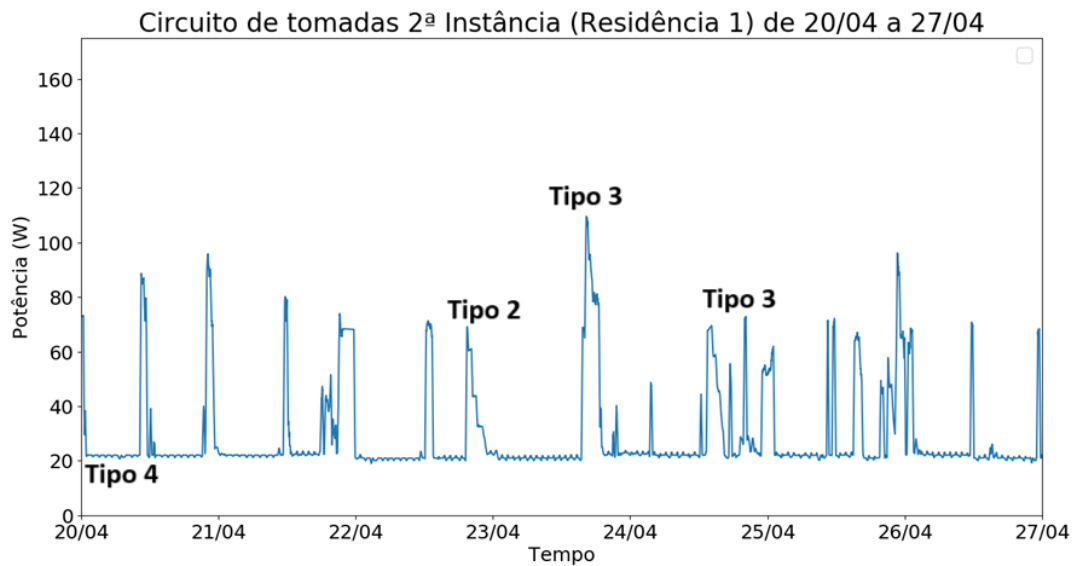
3.1.2.4 Circuito de Tomadas (2ª Instância)

Esta “carga”, ou melhor, este circuito, é composto por alguns aparelhos com assinaturas elétricas distintas, ligados em paralelo e registrados por um único conjunto de sensores. Apesar da falta de informações específicas sobre os aparelhos dentro deste circuito pelo banco de dados, é possível inferir as características das cargas através da análise de suas curvas de potência, ilustradas na Figura 12.

3.1.2.5 Lava-louças

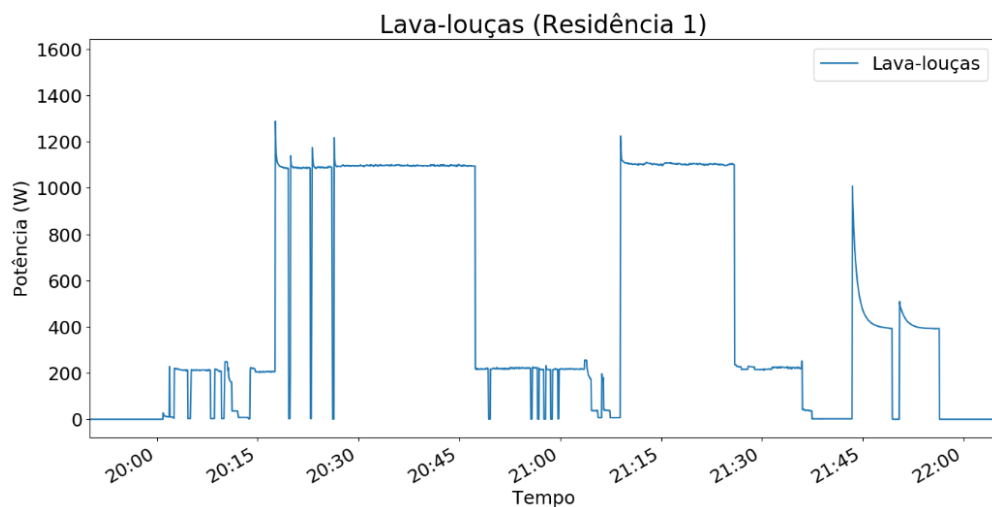
Esta carga apresenta uma assinatura elétrica claramente definida, ilustrada na Figura 13. A cada ativação, manteve-se consistente em seu comportamento, revelando um padrão de lavagem uniforme. Os horários de operação também foram consistentes, tendendo a concentrar-se no período noturno. Dada essa regularidade, fica claro que a carga se enquadra de forma precisa na categoria de estados finitos.

Figura 12 — Consumo de potência ativa do circuito de tomadas (2ª Instância) ao longo de uma semana (20/04 a 27/04/2011).



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 13 — Consumo de potência ativa da Lava-Louças.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

3.1.2.6 Micro-ondas

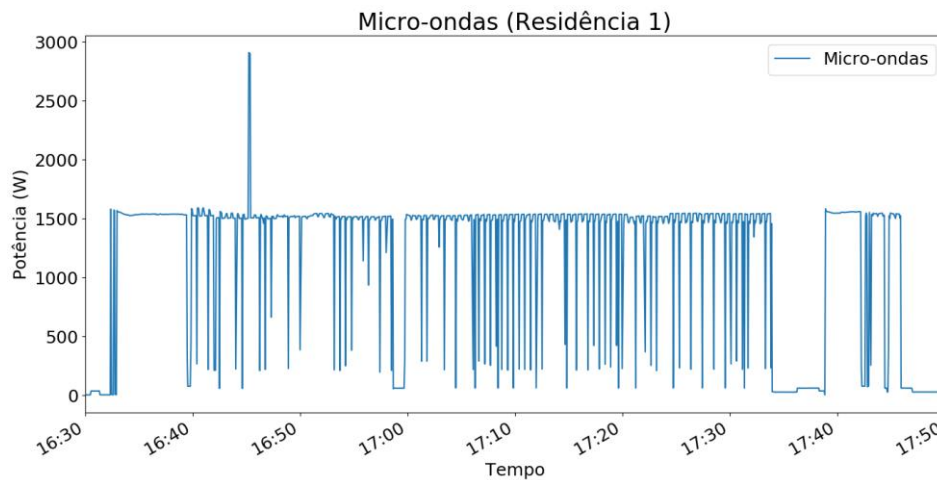
O micro-ondas, identificado como a sétima maior carga consumidora, distingue-se por sua elevada potência, aproximadamente 1500 W, e por sua assinatura elétrica única, conforme demonstrado no gráfico da Figura 14. O padrão de uso desse aparelho é caracterizado por eventos distintos e transições rápidas entre os estados de ligado e desligado, alinhando-o estreitamente com um modelo de carga binária.

Tabela 1 — Classificação das cargas monitoradas na Residência 1 do banco de dados REDD.

Cargas Monitoradas	Número de Cargas	Instâncias	Modelo de Carga	Perfil de Utilização	Potência média observada [W]
Iluminação	3	1	Estado Finito	Durante a noite	45 - 90
		2	Estado Finito	Várias vezes ao dia	65-100
		3	Estado Finito	Várias vezes ao dia	55
Geladeira	1	2	Estado Finito	Operação contínua	200 - 400
Aquecedor de Ambientes	1	2	-	Não utilizado	-
Tomadas	4	1	Variadas	Constantemente ligada	-
		2	Variadas	Constantemente ligada	-
		3	Ligado/Desligado	Raramente	1100
		4	Ligado/Desligado	Raramente	1100
Lavadora e secadora de Roupas	3	4	Estado Finito	Semanalmente	150
Lava-louças	1	2	Estado Finito	Algumas vezes na semana	500
Micro-ondas	1	2	Estado Finito	Várias vezes ao dia	120
Fogão Elétrico	1	2	-	Não utilizado	-
Forno Elétrico	1	2	Estado Finito	Utilizado poucas vezes	200 - 500
Cargas Desconhecidas	2	2	Estado Finito	Diariamente	100
		3	-	Não utilizado	-
	Total = 16				

Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 14 — Consumo de potência ativa do Micro-ondas.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

3.2 CENÁRIOS DE INVESTIGAÇÃO PROPOSTOS

Para aprofundar o entendimento sobre as dinâmicas da desagregação de cargas, propõem-se cinco cenários investigativos, que exploram as variáveis do número de cargas conhecidas, o número de cargas desconhecidas e o período de treinamento. A seleção desses cenários visa simular condições realistas, evidenciando como a expansão do número de cargas e do tempo de treinamento elevam os custos operacionais. Isso faz sentido ao se considerar que o treinamento demanda recursos como energia, equipamentos e pessoal qualificado, podendo, inclusive, interferir na rotina dos residentes. Outra dimensão que se pretende avaliar é o impacto da componente $e(t)$ da Equação 1, que representa um erro na modelagem do problema, analisando como a inclusão de cargas desconhecidas de baixo consumo como uma componente de erro influencia os resultados de desagregação.

A janela de desagregação é fixada em uma semana, de segunda a segunda (20/04/2011 a 27/04/2011), abrangendo dias úteis — quando a residência está usualmente desocupada durante o horário comercial, limitando o consumo a cargas como a geladeira e poucos outros dispositivos — e momentos de uso intensivo da rede elétrica. A duração da janela de desagregação varia conforme o cenário: para dois dias, entre 28 e 29/04, e para um dia, especificamente em 29/04. Estes períodos de treinamento foram escolhidos após análise minuciosa, identificando dias em que todos os aparelhos são utilizados pelo menos uma vez. A escolha de datas próximas

à janela de desagregação visa minimizar variações no funcionamento dos aparelhos devido a fatores externos, como alterações na temperatura ambiente ou defeitos. Esses dias foram selecionados após uma avaliação criteriosa que indicou momentos de uso de todos os aparelhos ao menos uma vez. A escolha de datas próximas à janela de desagregação visa minimizar variações no funcionamento dos aparelhos devido a fatores externos, como alterações na temperatura ambiente ou defeitos.

A seguir, apresenta-se uma síntese de cada cenário, elucidando os objetivos específicos de cada investigação e as expectativas em relação aos resultados.

3.2.1 *Cenário 1: 2D e TopK = 10*

No Cenário 1, o foco é examinar o impacto de um intervalo de treinamento de dois dias, durante o qual o algoritmo é instruído sobre os 10 maiores consumidores identificados na residência, conforme detalhado a Seção 3.1.2. Este experimento visa replicar uma situação em que os dados disponíveis sobre as cargas são restritos, mas, ainda assim, amplas informações sobre os estados e eventos de cada carga são fornecidas ao algoritmo durante o período de treinamento.

3.2.2 *Cenário 2: 2D e TopK = 13*

No Cenário 2, o objetivo é treinar o algoritmo dados referentes a todas as cargas que registraram um consumo superior a 2 kWh durante o mês. Esse limite de consumo foi estabelecido arbitrariamente para incluir todas as cargas que demonstram um nível de consumo significativo.

3.2.3 *Cenário 3: 2D e Todas as Cargas da Residência (TopK = 16)*

Considerando que o banco de dados disponibiliza os registros de consumo individual de todos os aparelhos instalados, este cenário se propõe a servir de base comparativa para os demais, ao adotar o período mais extenso de treinamento e incluir todas as cargas.

3.2.4 Cenário 4: 1D e TopK = 10

No Cenário 4, analisa-se o caso mais restritivo entre todos propostos, limitando os dados de treinamento às informações referentes às 10 maiores cargas em apenas um dia. Esta configuração busca avaliar os impactos de uma implementação com o mínimo de dados amostrais disponibilizados ao algoritmo para treinamento. O objetivo é explorar até que ponto a redução da quantidade de dados impacta na capacidade de desagregação do algoritmo e na precisão dos resultados obtidos.

3.2.5 Cenário 5: 1D e TopK = 13

No Cenário 5, introduz-se uma abordagem intermediária entre os demais cenários, avaliando os efeitos de treinar o algoritmo com uma amostra composta pelas 13 cargas mais relevantes (TopK = 13) em um período de treinamento de um dia.

3.3 TREINAMENTO DO ALGORITMO

A fase de treinamento do algoritmo é fundamental para determinar a qualidade da desagregação das cargas e a abordagem proposta visa explorar diferentes cenários com base nas maiores consumidoras de energia. Esta etapa é guiada por três perguntas-chave:

- 1) Qual o efeito da duração da janela de treinamento sobre o processo de desagregação?
- 2) É indispensável treinar o algoritmo com todas as cargas, ou é possível alcançar resultados adequados concentrando-se nas cargas de maior consumo?
- 3) Como a inclusão de um termo de erro afeta a precisão do sistema de desagregação?

A premissa de que um algoritmo deve ser treinado com dados que representem todos os estados operacionais de uma carga é amplamente aceita. No entanto, as análises das curvas de consumo das cargas e os dados da Tabela 1 sugerem a existência de variações significativas nos seus padrões de consumo. Assim, para assegurar a eficiência do algoritmo, é essencial examinar sua habilidade em identificar cargas frente a essas variações.

Portanto, a proposta é testar a reação do algoritmo CO a diferentes períodos de treinamento, oferecendo uma avaliação detalhada de sua performance e levando em consideração as limitações impostas pelo NILMTK, como a limitação no acesso a configurações específicas do algoritmo, incluindo ajustes de sensibilidade e na definição de estados transitórios e permanentes.

3.4 DESAGREGAÇÃO

Na etapa de desagregação, emprega-se o algoritmo para examinar o sinal de potência agregada da residência, requerendo a configuração da janela de observação estabelecida previamente, que abrange o período de 20/04/2011 a 27/04/2011.

3.5 MÉTODOS DE AVALIAÇÃO DOS RESULTADOS

Após finalizar o processo de desagregação, procede-se à análise dos resultados para aprofundar o entendimento do desempenho alcançado. O protocolo de avaliação sugerido inclui:

3.5.1 Avaliação Qualitativa: Uso de Gráficos

Para a avaliação qualitativa, os sinais desagregados são representados graficamente, sendo comparados com os sinais reais disponíveis no banco de dados. Esta comparação visual facilita a avaliação do desempenho do algoritmo CO frente às diversas cargas, permitindo identificar pontos fortes e fracos do algoritmo em diferentes dimensões.

3.5.2 Avaliação Quantitativa: F1-Score

Para a quantificação dos resultados, adota-se o F1-Score, que consolida as medidas de Precisão e Recall, englobando efetivamente os acertos e erros na identificação de TP, TN e FN, conforme apresentado nas Equações 4, 5 e 6 da Seção 2.9. Esta métrica está integrada ao NILMTK no módulo “metrics” e, como mostrado na Revisão Bibliográfica (Seção 2.9), é reconhecidamente uma das mais eficazes na avaliação de algoritmos de desagregação baseados em eventos.

3.6 CONCLUSÃO

Este capítulo apresenta a metodologia proposta para simulação e análise da técnica de desagregação de sinais de potência ativa em ambiente residencial, utilizando o *software* NILMTK versão 0.4.3, com ênfase nos módulos “*CombinatorialOptimization*” e “*Metrics*”, e o *software* Jupyter Notebook como IDE. A seleção dessas ferramentas baseia-se na sua importância para a área, conforme discutido anteriormente. A escolha de dados reais, do banco de dados REDD, reflete o objetivo de garantir que a avaliação do algoritmo de desagregação CO seja precisa e relevante.

Através da definição do conjunto de dados e a escolha da Residência 1 para estudo, este trabalho enfoca na análise de cargas com padrões de consumo significativos, estabelecendo uma base para a exploração dos cenários propostos. Esses cenários, variando no número de cargas conhecidas e na duração do treinamento, servem para testar a eficácia do algoritmo CO sob diversas condições. É importante enfatizar que o propósito deste estudo não é otimizar o algoritmo CO, mas sim aplicá-lo conforme a versão disponível no NILMTK, mantendo os parâmetros originais.

Durante a fase de treinamento, o algoritmo recebe dados detalhados de cada aparelho para construir um modelo de assinaturas elétricas. Na fase de desagregação, esses modelos são utilizados para analisar o sinal de potência agregada, observando-se um período específico de uma semana em todos os cenários.

Os métodos de avaliação dos resultados incluem análises qualitativas, por meio de comparações gráficas com os dados reais, e quantitativas, utilizando o F1-Score. Esta metodologia fornece uma análise aprofundada do desempenho do algoritmo CO, explorando como as diferentes configurações impactam na eficácia da desagregação.

Para aprofundamento técnico e referência ao código-fonte utilizado neste trabalho, o programa desenvolvido está disponível no Anexo A.

O próximo capítulo apresenta os dados obtidos a partir da aplicação da metodologia descrita.

4 RESULTADOS

Esta seção expõe os resultados alcançados ao aplicar os diferentes cenários de desagregação propostos previamente. Os cenários foram projetados para explorar a influência do número de cargas selecionadas e do período de treinamento na eficácia da técnica de desagregação. A análise cobre uma semana completa, abrangendo diferentes condições de uso dos aparelhos.

4.1 DESAGREGAÇÃO DAS CARGAS DE MAIOR CONSUMO

Os testes executados contemplam cinco cenários distintos, detalhados a seguir:

1º Cenário: Treinamento e desagregação das 10 principais cargas por consumo no mês (TopK 10), com um período de treino de dois dias.

2º Cenário: Treinamento e desagregação de todas as cargas com consumo superior a 1 kWh no mês (TopK 13), com um período de treino de dois dias.

3º Cenário: Treinamento e desagregação de todas as cargas instaladas, com um período de treino de dois dias.

4º Cenário: Similar ao primeiro cenário, mas com a janela de treinamento ajustada para um dia.

5º Cenário: Similar ao segundo cenário, mas com a janela de treinamento ajustada para um dia.

Para cada cenário, o algoritmo foi inicialmente treinado com dados individuais de cada carga durante o período designado para o treinamento, seguido pelo processo de desagregação. A biblioteca *Matplotlib* é utilizada para geração dos gráficos. Os resultados são mostrados nas próximas subseções.

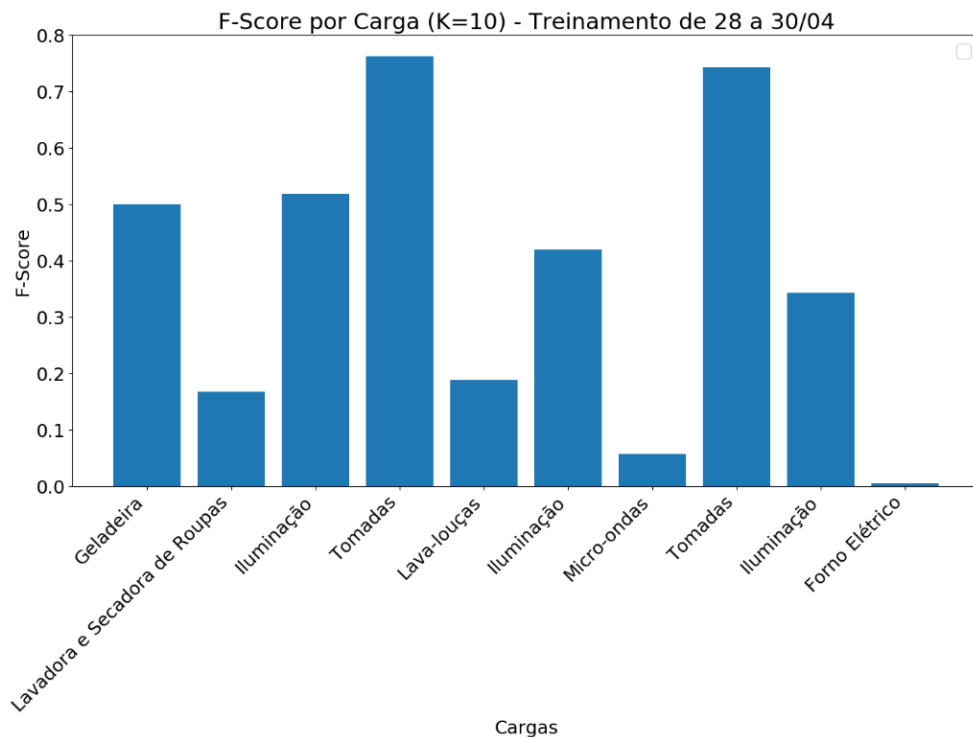
4.1.1 Janela de Treinamento de Dois Dias (2D)

Nesta subseção serão mostrados os resultados de desagregação para os três primeiros cenários.

4.1.1.1 Cenário 1: 2D e TopK = 10

O primeiro cenário testado envolve a desagregação das 10 cargas que mais consumiram em um mês de registros. A Figura 15 apresenta o F1-Score alcançado.

Figura 15 — F1-Score do cenário de desagregação das TopK 10 cargas com a janela de treinamento de dois dias.

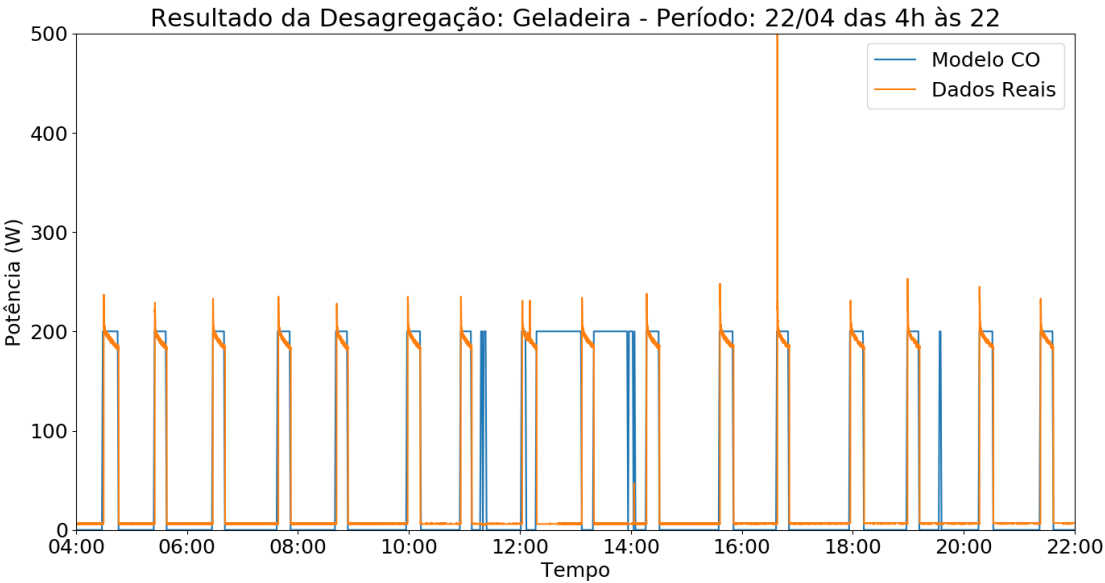


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

A Figura 16 ilustra o sinal de potência resultante da desagregação da geladeira sobreposto ao sinal agregado real. A Figura 17 exibe a mesma janela de tempo da Figura 16, mas destaca as limitações do algoritmo em diferenciar eventos da geladeira em meio à atividade simultânea de múltiplas cargas.

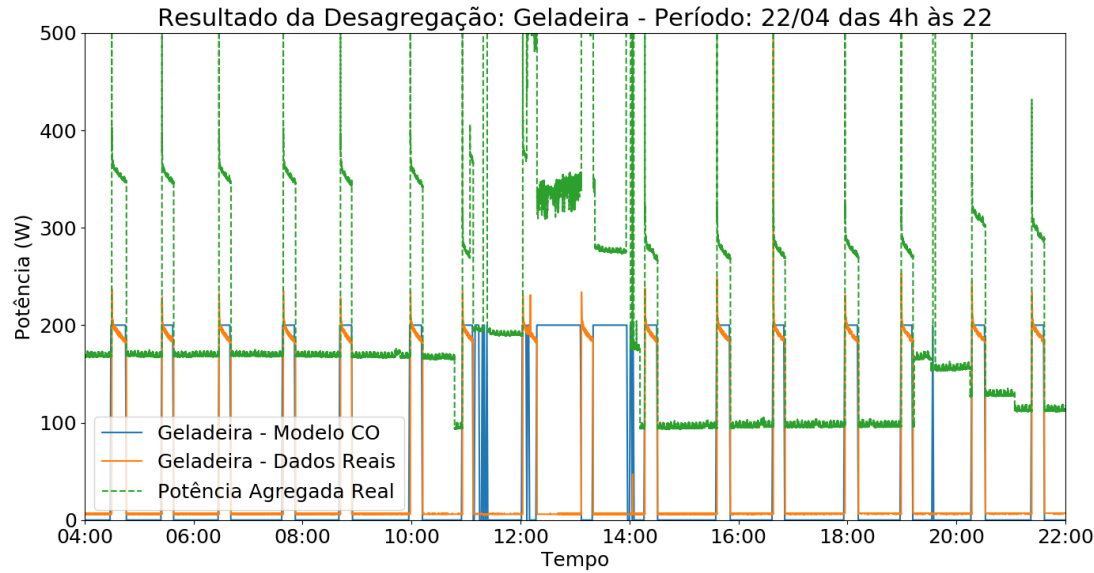
A Figura 18 ilustra o consumo real das cargas detalhando um dos ciclos de operação da geladeira mostrados na Figura 16. Neste período houve pouco consumo das demais cargas da residência. Já a Figura 19 ilustra o resultado da desagregação no mesmo momento do sinal exibido na Figura 18. A Figura 19 destaca erros do algoritmo na desagregação da geladeira, atribuindo seu consumo a outros aparelhos.

Figura 16 — Resultado da desagregação da geladeira durante aproximadamente um dia (1º Cenário).



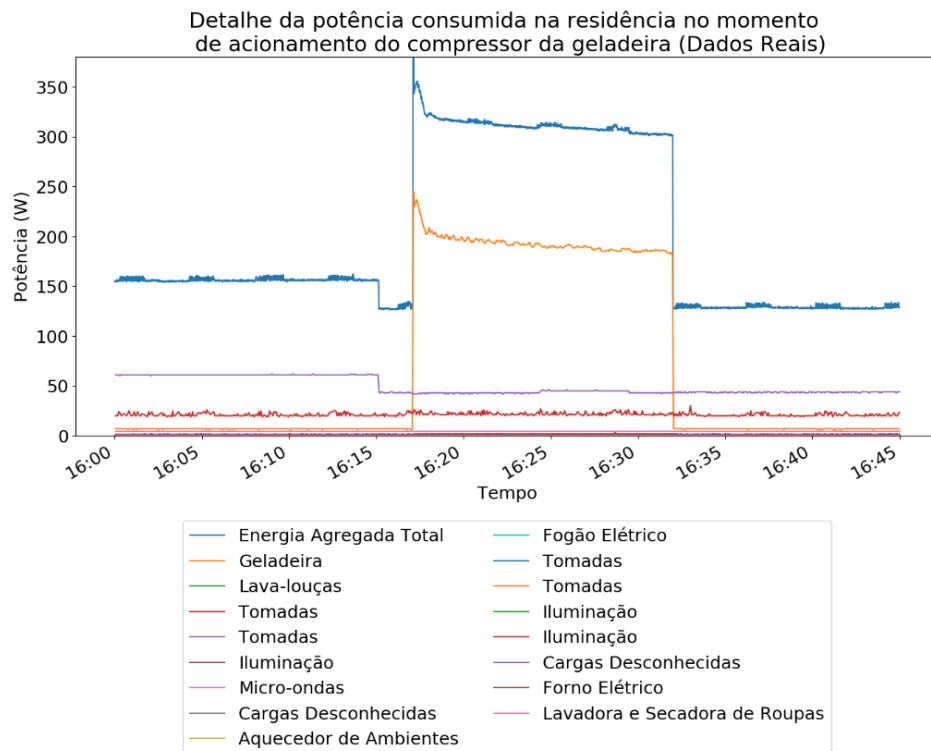
Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 17 — Resultado da desagregação da geladeira no mesmo período mostrado na Figura 16.



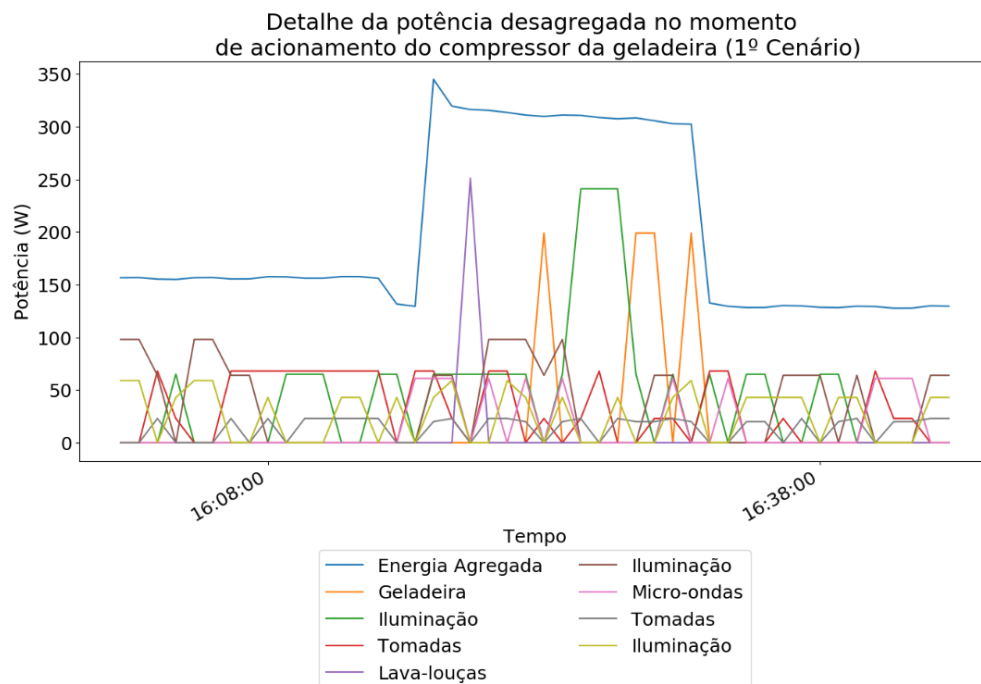
Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 18 — Detalhe do momento de acionamento do compressor da geladeira, com sobreposição do consumo individual e agregado cargas.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 19 — Detalhe do resultado da desagregação em um ciclo de acionamento da geladeira, evidenciando erros no processo (1º Cenário).

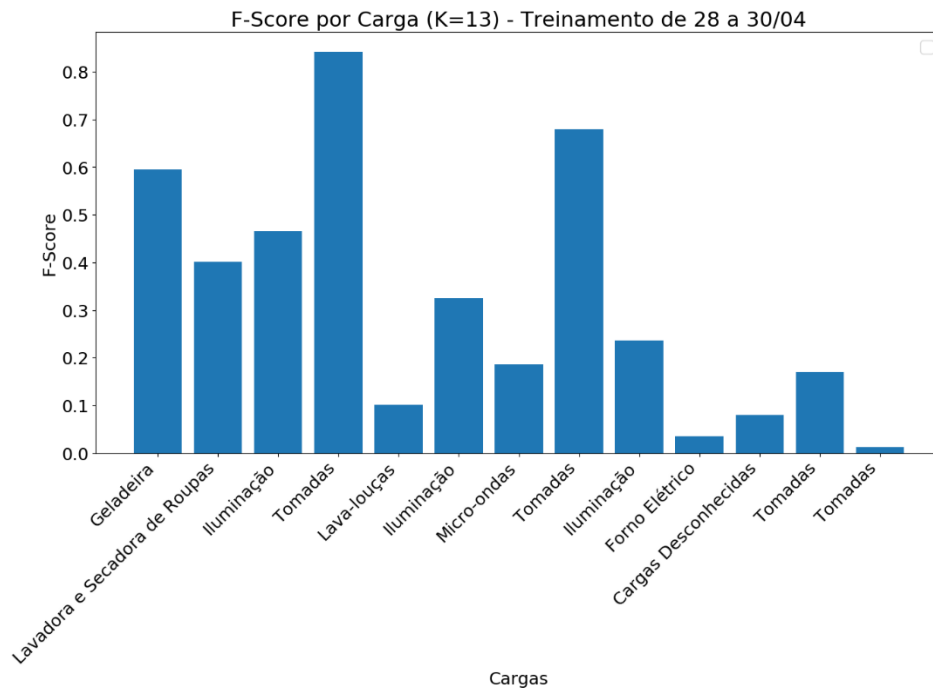


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

4.1.1.2 Cenário 2: 2D e TopK = 13

Neste segundo cenário são todas as cargas que apresentaram consumo representativo são informadas ao algoritmo durante o treinamento. A Figura 20 mostra o F1-Score resultante.

Figura 20 — F1-Score da desagregação das TopK 13 cargas com a janela de treinamento de dois dias.



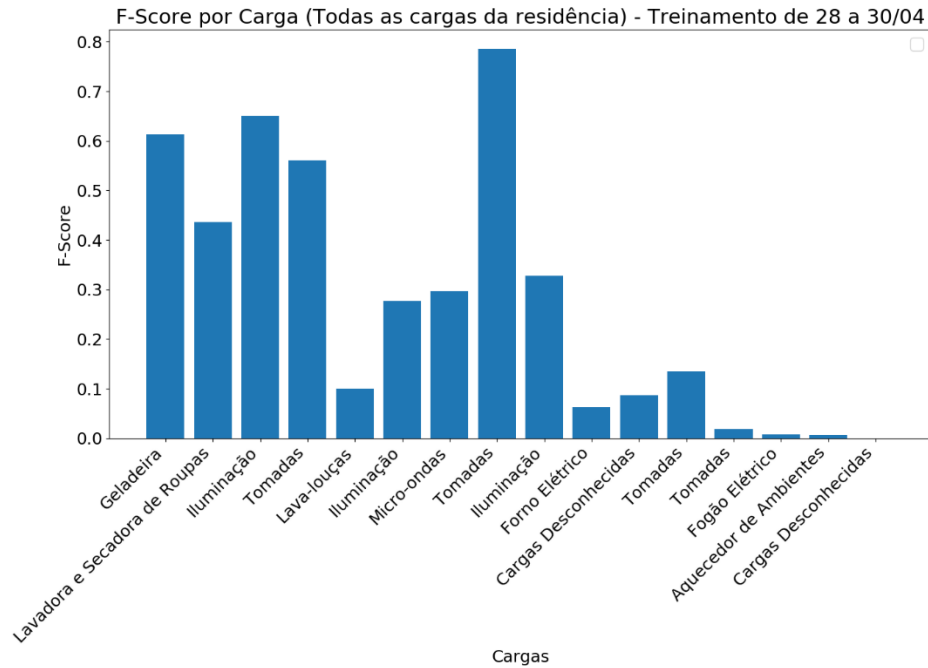
Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

4.1.1.3 Cenário 3: 2D e Todas as Cargas da Residência (TopK 16)

No terceiro cenário todas as cargas instaladas são ensinadas ao algoritmo. A Figura 21 mostra o F-Score obtido.

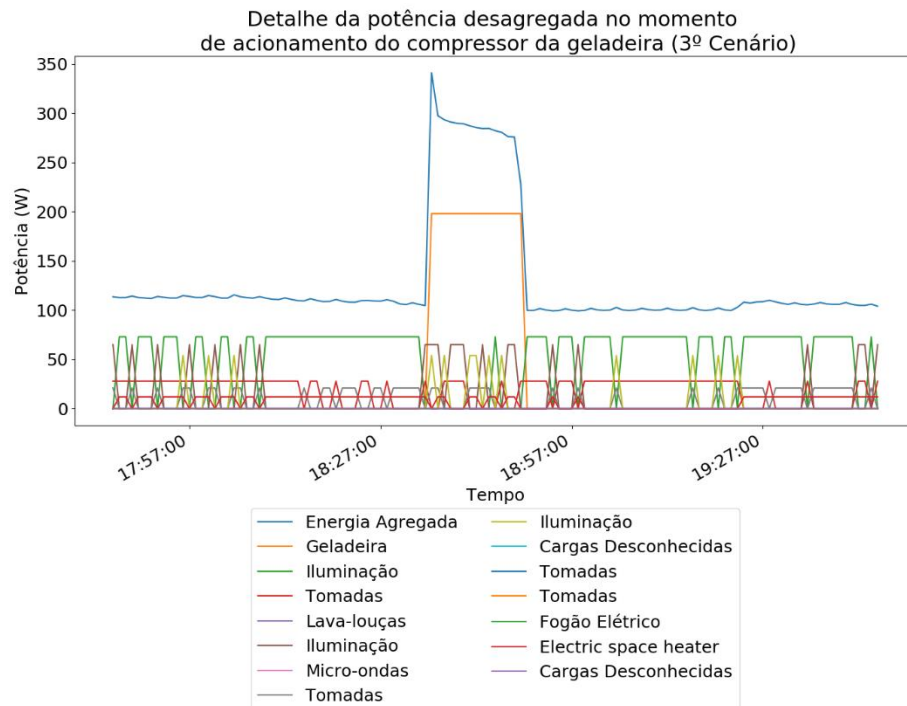
A Figura 22 ilustra um resultado bem-sucedido na desagregação da geladeira. Ainda assim é possível notar erros na desagregação das outras cargas.

Figura 21 — F-Score do cenário de desagregação de todas as cargas instaladas na residência com a janela de treinamento de dois dias.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

Figura 22 — Detalhe do resultado da desagregação em um dos ciclos de acionamento da geladeira (3º Cenário).



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

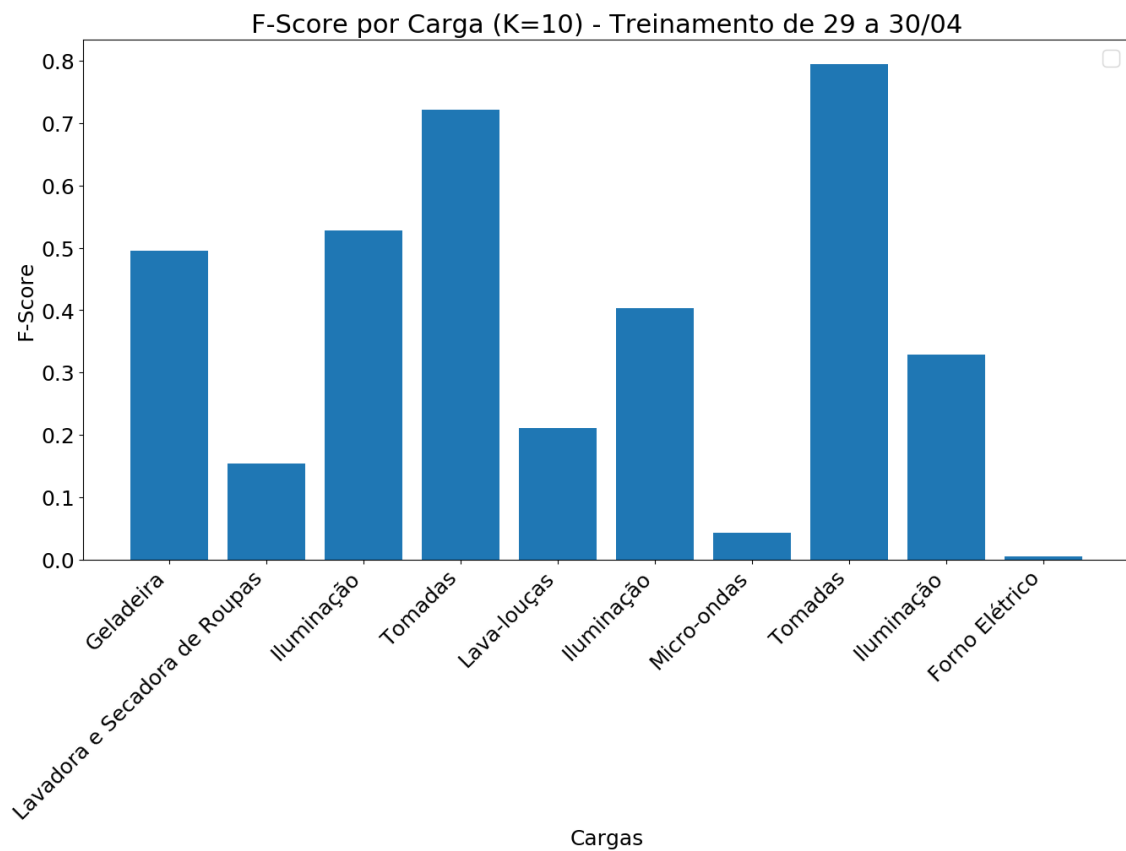
4.1.2 Janela de Treinamento de Um Dia

Esta subseção expõe os resultados da desagregação com a janela de treinamento definida para um único dia, aplicando a mesma janela de desagregação dos cenários anteriores. Explora-se a variação no número de cargas treinadas, alternando entre as 10 maiores consumidoras (TopK 10) e as 13 maiores (TopK 13).

4.1.2.1 Cenário 4: 1D e TopK = 10

Neste cenário é feita a desagregação das TopK 10 cargas com a janela de treinamento de um dia. A Figura 23 mostra a métrica F1-Score obtida considerando os resultados de desagregação.

Figura 23 — F1-Score do cenário de desagregação das TopK 10 cargas com a janela de treinamento de um dia.

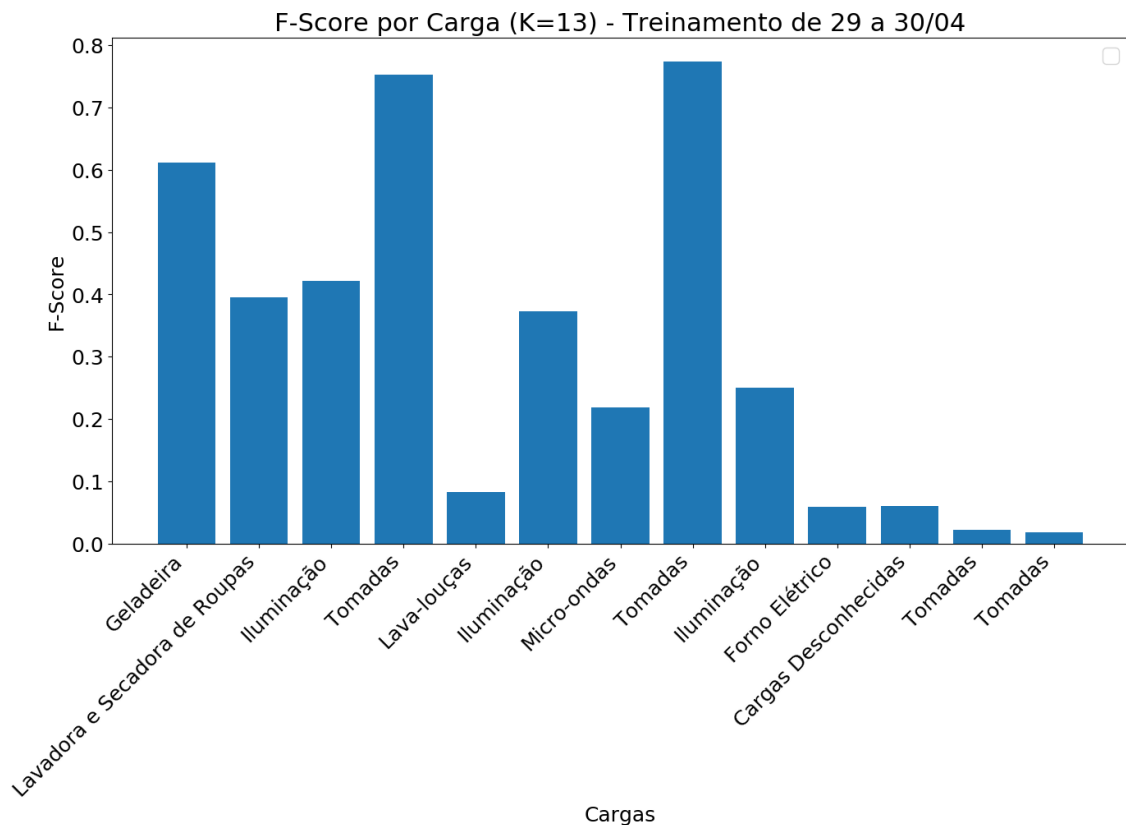


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

4.1.2.2 Cenário 5: 1D e TopK = 13

Neste cenário é feita a desagregação das TopK 13 cargas com a janela de treinamento de um dia. A Figura 24 mostra a métrica F1-Score obtida considerando os resultados de desagregação.

Figura 24 — F1-Score do cenário de desagregação das TopK 13 cargas com a janela de treinamento de um dia.

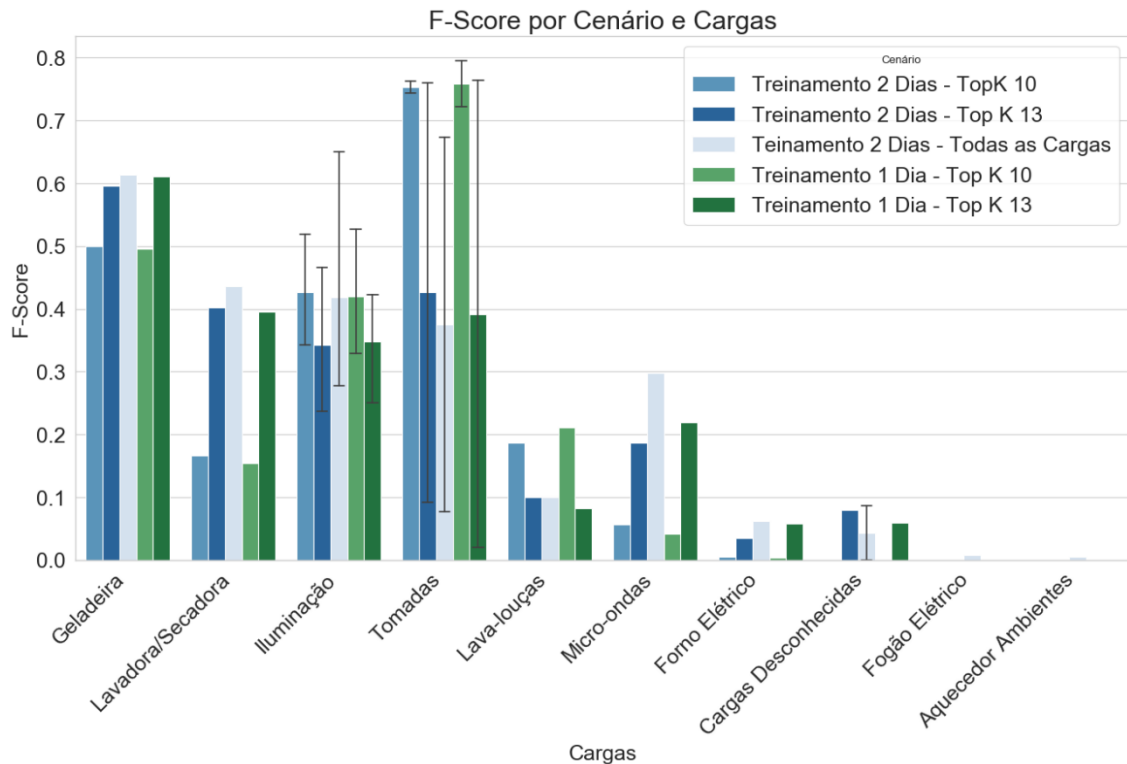


Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

4.2 COMPARAÇÃO DOS RESULTADOS

A Figura 25 mostra uma comparação dos resultados da métrica F1-Score entre todos os cenários testados. A análise das barras de cada aparelho deve ser feita da esquerda para a direita, indicando como o aumento no número de cargas monitoradas impactou o resultado para cada carga. As cores azuis indicam os cenários com janela de treinamento de 2 dias e os dados em cor verde indicam os cenários com um dia de treinamento.

Figura 25 — Comparação da métrica F1-Score para todos os cenários testados.



Fonte: Adaptado a partir do banco de dados REDD (KOLTER, JOHNSON, 2011) pelo Autor.

4.3 DISCUSSÃO

A análise dos resultados revela uma relação complexa entre a janela de treinamento, a inclusão de mais cargas e o desempenho na desagregação das cargas elétricas. O tipo e perfil de uso das cargas também teve influência direta com os resultados. Neste sentido, a discussão pode ser feita para algumas tendências observadas.

4.3.1 Consolidação do Aprendizado com Janelas de Treinamento Reduzidas

A modificação da janela temporal de treinamento visou reforçar o aprendizado do algoritmo ao fornecer mais assinaturas elétricas nesta etapa. Interessantemente, a expansão da janela temporal resultou em uma estabilização dos resultados para cenários equiparados em número de cargas sugerindo que um intervalo reduzido de aprendizado é suficiente para a assimilação das assinaturas elétricas, inclusive para cargas menos frequentes. (Figura 25, cargas: Geladeira, Lavadora e Secadora,

Iluminação, Tomadas, Lava-louças, Micro-ondas, Forno Elétrico e até mesmo as Cargas Desconhecidas).

Cargas cujo desempenho decai com a ampliação da janela de treinamento representam exceções, cuja análise sugere a influência de fatores externos ao método CO. Ao interpretar circuitos contendo múltiplas cargas com assinaturas distintas como uma carga única, notadamente os circuitos de iluminação (Figura 11) e tomadas (Figura 12), o treinamento do algoritmo associa a esta “carga” vários estados com níveis de potência variados, espalhando seus *clusters* pelo espaço de assinaturas. Aumentar o tempo significa que o algoritmo consolida estes *clusters* sobre regiões de potência ocupadas por outras cargas, conduzindo a erros de classificação.

4.3.2 Desempenho Comprometido por Assinaturas Elétricas Similares

Conforme antecipado, cargas com assinaturas elétricas semelhantes enfrentam maiores desafios na desagregação. Esta situação é observada nos circuitos de iluminação e na lava-louças, onde a competição por regiões no espaço de assinaturas e a sobreposição com outras cargas exacerbaram os desafios de desagregação. A primeira carga ocupa uma região de estados na faixa de 45 W a 100 W, que concorre com os circuitos de tomadas. Já a última compartilha uma região de potência próxima da geladeira. Ampliar o número de cargas conhecidas pelo algoritmo é, nestes casos, a estratégia mais eficaz para aprimorar o desempenho do modelo e é apresentada na subseção a seguir.

4.3.3 Acréscimo de Cargas Conhecidas e Melhoria de Desempenho

A comparação entre diversos cenários das cargas geladeira, lavadora e secadora e micro-ondas revelou uma tendência clara: o enriquecimento do banco de dados do modelo com uma maior variedade de cargas resultou em melhorias superiores a 0,1 no desempenho, conforme evidenciado pelo aumento nos valores de F-Score. Este resultado sugere que aumentar o número de estados conhecidos diminui o tamanho dos *clusters* e melhor representa as assinaturas elétricas de cada aparelho, facilitando a correta atribuição de eventos durante a desagregação.

4.3.4 Análise Específica: A Geladeira

Dentre as cargas analisadas, a geladeira, com o maior consumo energético registrado (44 kWh/mês), apresentou um dos melhores desempenhos em termos de F-Score em todos os cenários, variando de aproximadamente 0,5 a 0,6. Tal resultado sugere que a regularidade cíclica e estável fornece informações robustas o suficiente para obter sucesso em sua identificação mesmo com intervalos de treinamento reduzidos. Este resultado destaca a eficácia da metodologia em capturar e interpretar padrões de consumo estáveis. Entretanto, a análise detalhada do consumo e dos sinais estimados destaca desafios na reconstrução precisa dos sinais de cada carga, evidenciando tanto a capacidade quanto as limitações do método utilizado.

Os resultados mostrados nas Figuras 16 a 19, proporcionam uma análise comparativa entre os consumos real e estimados em momentos de baixo e alto consumo. Ao examinar estes resultados fica evidente que a técnica apresenta desafios significativos na reconstrução dos sinais individuais de cada carga, especialmente durante picos de demanda. Mesmo em situações em que o processo de desagregação da geladeira obteve êxito (Figura 22), é possível notar a identificação de outros eventos inexistentes e a atribuição de consumo residual a aparelhos que não estavam presentes naquele momento. Contudo, é digno de nota que, apesar dessas dificuldades, o algoritmo ainda assim demonstrou acurácia na identificação dos eventos de acionamento e desacionamento da geladeira mesmo quando desconhecia vários aparelhos na rede.

4.3.5 Observações Sobre o Funcionamento do Algoritmo

A avaliação das Figuras 18, 19 e 22, proporciona uma visão detalhada das técnicas empregadas pelo algoritmo e suas consequências para a precisão e eficácia do modelo. É possível notar o emprego de *downsampling*, técnica apresentada no capítulo 2.5, que suaviza os sinais agregados antes do processo de desagregação. Esta abordagem está intrinsecamente ligada a implementação do algoritmo no NILMTK, cuja configuração padrão opera os sinais agregados a uma resolução máxima de 60 segundos, o que, embora reduza a complexidade computacional, impõe limitações significativas na capacidade do algoritmo de discernir eventos de curta duração. Entretanto, o uso dessa técnica também traz vantagens ao diminuir sua

sensibilidade a transitórios, o que de fato é proposto pelo criador do método (HART, 1992). Outra característica é possível ser observada na prática: a aplicação de heurísticas específicas que funcionam como filtros para regular a detecção de eventos próximos no tempo, apresentadas na Seção 2.7.3, que também são implementadas com o intuito de diminuir a ocorrência de erros de classificação.

4.3.6 Uma Breve Discussão dos Impactos do Erro $e(t)$

A comparação dos resultados obtidos nos cenários 1, 2 e 3 mostra que o desempenho do algoritmo decai ao operar num cenário onde não se conhecem todas as cargas. A média geral dos F1-Scores em cada cenário é:

Cenário 1: 0,427414

Cenário 2: 0,481252

Cenário 3: 0,472228

De acordo com a Equação 1, o termo indicativo de erro $e(t)$ pode significar, entre outras fontes de erro, a presença de outros aparelhos não informados ao algoritmo na mesma rede elétrica. Nesse sentido, há um ganho inicial de desempenho de 12,59 % ao adicionar 3 aparelhos conhecidos. No entanto, ao treiná-lo com todas as 16 cargas há uma relativa estabilização do desempenho. Isso mostra que pequenas cargas podem ser inseridas na rede sem impacto no resultado geral. De certa forma isso mostra algum nível de flexibilidade do algoritmo. Ainda assim é prudente afirmar que são necessários mais estudos para entender o comportamento do algoritmo nesses cenários.

4.3.7 Aquecedor de Ambientes e Cargas Desconhecidas

O aquecedor de ambientes não foi utilizado durante o período observado, justificando seu baixo desempenho geral. Já a primeira instância das cargas desconhecidas é a única que apresenta consumo entre as duas cargas desconhecidas monitoradas. Seus estados circulam entre 6W e 12 W, sugerindo serem carregadores ou fontes externas. Seu baixo desempenho em todos os cenários sugere a dificuldade do algoritmo em desagregar pequenas cargas.

5 CONCLUSÕES, SUGESTÕES DE TRABALHOS FUTUROS E CONSIDERAÇÕES FINAIS

Este trabalho proporcionou um estudo aprofundado sobre a metodologia NILM, destacando suas características e a inter-relação entre suas etapas para alcançar o objetivo principal de investigar e avaliar o desempenho do algoritmo CO na desagregação de cargas elétricas, utilizando o *software* NILMTK. A investigação incluiu a aplicação e a avaliação da técnica de desagregação de sinais elétricos, empregando dados de consumo energético de uma residência real, disponibilizados pelo banco de dados REDD. O estudo enfrentou desafios significativos, como a complexidade de aprendizado do NILMTK, a ausência de um banco de dados ideal e, especialmente, a existência de mais perguntas do que respostas nessa área de pesquisa teorizada a pouco mais de 30 anos.

O estudo enfatizou a importância das assinaturas elétricas, que englobam diversas características com distintos níveis de complexidade, na implementação bem-sucedida de aplicações NILM. Este trabalho destaca a importância de avaliar minuciosamente essas características, reconhecendo a existência de assinaturas elétricas, muitas vezes simples, capazes de descrever e permitir a identificação de cargas específicas na rede elétrica, tais como frequência de utilização, hora do dia e clima.

A principal conclusão deste trabalho ressalta a inviabilidade de desenvolver um sistema NILM universal, devido à diversidade de cargas e suas assinaturas elétricas únicas. O que se pode inferir é que cada tipo de carga demanda uma abordagem distinta, e algumas características das assinaturas podem ser extrapoladas para cargas com propriedades similares. Adicionalmente, este estudo conclui que o projeto e avaliação da eficácia de uma solução NILM não pode ser fragmentada, requerendo a consideração do conjunto total de cargas.

Os resultados indicam que o maior impacto na desagregação ocorreu com o aumento do número de cargas conhecidas pelo algoritmo CO, o que já era esperado e sugere que para um bom desempenho deste método, é necessário que o algoritmo conheça as cargas em sua integralidade. De maneira contrária, variações no tempo de treinamento apresentaram impacto marginal no desempenho geral, sugerindo que períodos de aprendizado breves são suficientes para a assimilação das assinaturas elétricas, mesmo para cargas de uso esporádico. Este resultado aponta para a

necessidade de estudos futuros focados na influência do tempo de treinamento e do conhecimento das cargas no processo de desagregação. As limitações deste estudo ao utilizar um conjunto de dados real, resultam no atendimento parcial do objetivo específico de “Estabelecer correlações entre os resultados da desagregação, as assinaturas elétricas e os cenários propostos, buscando inferir conclusões aplicáveis a outros contextos e identificar os principais fatores que influenciam o desempenho do algoritmo CO”, pois não se atingiu o controle de todas as variáveis que compõe o sinal agregado analisado. Estes pontos ainda merecem a atenção em trabalhos futuros.

Em particular, a análise do desempenho da geladeira demonstrou a capacidade do algoritmo em aprender profundamente suas assinaturas elétricas, resultando em um desempenho destacado. Esta observação sugere potencial para a detecção de cargas de Estado Finito, recomendando-se a diminuição da sensibilidade do algoritmo na detecção de eventos para melhorar os resultados.

Destaca-se, ainda, uma limitação fundamental nesta implementação: cargas com assinaturas elétricas similares induzem o algoritmo CO a erros. Nestes casos recomenda-se a exploração de abordagens alternativas na metodologia NILM, como o aumento do número de assinaturas observadas e a adoção de técnicas não baseadas em eventos, capazes de lidar de forma mais eficaz com variações contínuas (IQBAL, MALIK, *et al.*, 2021).

Além disso, a inclusão de mais cargas e o aumento da janela de treinamento afetaram o tempo de processamento e elevaram o custo computacional e a complexidade do modelo. Diante do objetivo específico de "Identificar cenários reais de consumo de energia nos quais a metodologia NILM baseada em eventos possa trazer benefícios significativos", verifica-se que setores como frigoríficos, supermercados e avicultura, que dependem de cargas críticas como refrigeradores e congeladores, poderiam se beneficiar de soluções inspiradas neste estudo. Nessas aplicações, a prioridade é assegurar o funcionamento adequado desses equipamentos em momentos específicos. Uma implementação efetiva dessa metodologia requereria, como requisito mínimo, a capacidade de verificar a ocorrência de eventos ligados às cargas críticas, sem a necessidade de um monitoramento contínuo do consumo de energia. É importante destacar que cargas críticas acionadas por relés mecânicos, ainda muito utilizados, podem apresentar falhas sendo que, em

última medida, a única forma de garantir seu estado de operação é observando seu consumo.

5.1 SUGESTÕES DE TRABALHOS FUTUROS

Várias dúvidas persistem sem resposta acerca da influência que a quantidade de eventos informados e o número de cargas conhecidas exercem sobre os resultados. Recomenda-se que pesquisas futuras abordem essas variáveis de forma mais controlada, visando elucidar suas relações e impactos de maneira precisa.

No âmbito da otimização do desempenho do algoritmo, propõe-se a avaliação dos efeitos decorrentes da variação dos parâmetros associados ao *downsampling*. Além disso, a exploração da implementação de outras técnicas de pré-processamento dos sinais merece atenção especial, especialmente aquelas que incluem a compensação dos sinais de potência em resposta às flutuações na tensão de fornecimento. Estas estratégias apresentam potencial de aprimorar a robustez das soluções NILM frente às variações ambientais e operacionais.

5.2 CONSIDERAÇÕES FINAIS

Em reconhecimento à extraordinária visão do autor do artigo (HART, 1992), este trabalho empenhou-se na exploração dos seus fundamentos, visando contribuir para o contínuo avanço no campo da desagregação energética. Envolvendo-se na pesquisa da área NILM, o autor se deparou com desafios que destacaram a necessidade de um conhecimento abrangente em diversas disciplinas, ultrapassando barreiras tradicionais ao envolver conceitos de engenharia elétrica, comunicações e ciência de dados. Atualmente, as principais universidades e empresas em todo mundo concentram esforços no desenvolvimento de sistemas similares, enfrentando desafios análogos aos abordados neste trabalho. Uma compreensão mais precisa do consumo energético de pessoas e equipamentos abre caminho para propostas de soluções de consumo mais eficientes e inteligentes.

O potencial transformador inerente a esta área é imenso e ferramentas como NILM apresentam-se como aliadas na busca por soluções para os grandes desafios enfrentados pela humanidade. Este trabalho contribuiu para evidenciar que a tecnologia NILM confere valor a dados antes considerados inúteis. Os sinais de

consumo elétrico escondem informações preciosas e a tecnologia NILM possibilita o monitoramento de diversas cargas a um custo significativamente inferior ao da ILM, promovendo uma alternativa viável e sustentável.

Ao abordar questões pertinentes à implementação de sistemas NILM baseados em eventos, este trabalho reconhece a complexidade inerente a esses sistemas. As análises realizadas examinaram a influência de uma variedade de fatores variáveis, oferecendo respostas perspicazes que, apesar de suas limitações, têm potencial para serem aplicadas em diversos contextos. Desafios como o número e tipo de cargas avaliadas, a duração do período de treinamento, a localização do monitoramento e as incertezas de medição representam desafios complexos de serem totalmente controlados em cenários reais.

Ao concluir a análise, este estudo não busca esgotar as complexidades da área, mas sim incentivar sua contínua exploração. Este estudo oferece uma contribuição significativa ao avanço do conhecimento e das práticas de desagregação energética, indicando claramente que muitos questionamentos permanecem e novos surgirão à medida que esta área dinâmica evolui.

REFERÊNCIAS

ABUBAKAR, I., KHALID, S. N., MUSTAFA, M. W., *et al.* "An overview of Non-intrusive load monitoring methodologies", **2015 IEEE Conference on Energy Conversion, CENCON 2015**, p. 54–59, 2015. DOI: 10.1109/CENCON.2015.7409513. . Acesso em: 26 fev. 2024.

ANDERSON, K. D., OCNEANU, A., CARLSON, D. R., *et al.* "BLUED : A Fully Labeled Public Dataset for Event-Based Non-Intrusive Load Monitoring Research", **Workshop on Data Mining Applications in Sustainability (SustKDD)**, 2012. . Acesso em: 30 nov. 2023.

ANGELIS, G. F., TIMPLALEXIS, C., KRINIDIS, S., *et al.* "NILM applications: Literature review of learning approaches, recent developments and challenges", **Energy and Buildings**, v. 261, p. 111951, 15 abr. 2022. DOI: 10.1016/J.ENBUILD.2022.111951. . Acesso em: 1 mar. 2024.

BANDEIRA, P., MARTINS, M. **NON-INTRUSIVE INDUSTRIAL LOAD MONITORING ON A FACTORY IN BRAZIL**. 2020. COPPE - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2020. . Acesso em: 4 dez. 2023.

BATRA, N., KELLY, J., PARSON, O., *et al.* "NILMTK: An Open Source Toolkit for Non-intrusive Load Monitoring", **e-Energy 2014 - Proceedings of the 5th ACM International Conference on Future Energy Systems**, p. 265–276, 15 abr. 2014. DOI: 10.1145/2602044.2602051. Disponível em: <http://arxiv.org/abs/1404.3878>. Acesso em: 20 nov. 2023.

BATRA, N., PARSON, O., BERGES, M., *et al.* "A comparison of non-intrusive load monitoring methods for commercial and residential buildings", 27 ago. 2014. Disponível em: <https://arxiv.org/abs/1408.6595v1>. Acesso em: 20 nov. 2023.

BERGES, M. E., GOLDMAN, E., MATTHEWS, H. S., *et al.* "Enhancing electricity audits in residential buildings with nonintrusive load monitoring", **Journal of Industrial Ecology**, v. 14, n. 5, p. 844–858, out. 2010. DOI: 10.1111/J.1530-9290.2010.00280.X. . Acesso em: 2 mar. 2024.

BRASIL. "ANUÁRIO ESTATÍSTICO DE ENERGIA ELÉTRICA 2022 ANUÁRIO 2022". , 2022. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/anuario-estatistico-de-energia-eletrica>. Acesso em: 24 nov. 2023.

BRASIL. **Pesquisa de Posse e Hábitos de Uso de Equipamentos Elétricos na Classe Residencial 2019**. 2019. Disponível em: <https://q.eletrabras.com/pt/Paginas/PPH-2019.aspx>. Acesso em: 24 nov. 2023.

BRASIL. "Procedimentos de Distribuição de Energia Elétrica no Sistema Elétrico Nacional – PRODIST", **Agência Nacional de Energia Elétrica**, 2015. . Acesso em: 20 nov. 2023.

BUCCI, G., CIANCETTA, F., FIORUCCI, E., *et al.* "State of art overview of Non-Intrusive Load Monitoring applications in smart grids", **Measurement: Sensors**, v. 18, p. 100145, 1 dez. 2021. DOI: 10.1016/J.MEASEN.2021.100145. . Acesso em: 1 mar. 2024.

CARRIE ARMEL, K., GUPTA, A., SHRIMALI, G., *et al.* "Is disaggregation the holy grail of energy efficiency? The case of electricity", **Energy Policy**, v. 52, p. 213–234, 1 jan. 2013. DOI: 10.1016/J.ENPOL.2012.08.062. . Acesso em: 21 nov. 2023.

"Evaluation of Advanced Metering System (AMS) Deployment in Texas Navigant Consulting (PI) LLC". 2010. Disponível em: www.navigantconsulting.com. Acesso em: 22 nov. 2023.

FAUSTINE, A., MVUNGI, N. H., KAIJAGE, S., *et al.* "A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem", 2 mar. 2017. Disponível em: <https://arxiv.org/abs/1703.00785v3>. Acesso em: 19 nov. 2023.

GOPINATH, R., KUMAR, M., PRAKASH CHANDRA JOSHUA, C., *et al.* "Energy management using non-intrusive load monitoring techniques – State-of-the-art and future research directions", **Sustainable Cities and Society**, v. 62, p. 102411, 1 nov. 2020. DOI: 10.1016/J.SCS.2020.102411. . Acesso em: 28 fev. 2024.

HAQ, A. U., JACOBSEN, H. A. "Prospects of Appliance-Level Load Monitoring in Off-the-Shelf Energy Monitors: A Technical Review", **Energies 2018, Vol. 11, Page 189**, v. 11, n. 1, p. 189, 12 jan. 2018. DOI: 10.3390/EN11010189. Disponível em: <https://www.mdpi.com/1996-1073/11/1/189/htm>. Acesso em: 21 nov. 2023.

HART, G. W. "Nonintrusive Appliance Load Monitoring", **Proceedings of the IEEE**, v. 80, n. 12, p. 1870–1891, 1992. DOI: 10.1109/5.192069. .

HOSSEINI, S. S., AGBOSSOU, K., KELOUWANI, S., *et al.* "Non-intrusive load monitoring through home energy management systems: A comprehensive review", **Renewable and Sustainable Energy Reviews**, v. 79, p. 1266–1274, 1 nov. 2017. DOI: 10.1016/J.RSER.2017.05.096. . Acesso em: 19 nov. 2023.

IQBAL, H. K., MALIK, F. H., MUHAMMAD, A., *et al.* "A critical review of state-of-the-art non-intrusive load monitoring datasets", **Electric Power Systems Research**, v. 192, p. 106921, 1 mar. 2021. DOI: 10.1016/J.EPSR.2020.106921. . Acesso em: 1 dez. 2023.

KASELIMI, M., PROTOPAPADAKIS, E., VOULODIMOS, A., *et al.* "Towards trustworthy Energy Disaggregation: A review of challenges, methods and perspectives for Non-Intrusive Load Monitoring", **Sensors**, v. 22, n. 15, 5 jul. 2022. DOI: 10.3390/s22155872. Disponível em: <https://arxiv.org/abs/2207.02009v1>. Acesso em: 27 fev. 2024.

KELLY, J., KNOTTENBELT, W. "The UK-DALE dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes", **Scientific Data** **2015 2:1**, v. 2, n. 1, p. 1–14, 31 mar. 2015. DOI: 10.1038/sdata.2015.7. Disponível em: <https://www.nature.com/articles/sdata20157>. Acesso em: 30 nov. 2023.

KHAZAEI, M., STANKOVIC, L., STANKOVIC, V. "Evaluation of low-complexity supervised and unsupervised NILM methods and pre-processing for detection of multistate white goods", **NILM 2020 - Proceedings of the 5th International Workshop on Non-Intrusive Load Monitoring**, p. 34–38, 18 nov. 2020. DOI: 10.1145/3427771.3427850. Disponível em: <https://dl.acm.org/doi/10.1145/3427771.3427850>. Acesso em: 25 fev. 2024.

KLEMENJAK, C., KOVATSCH, C., HEROLD, M., *et al.* "A synthetic energy dataset for non-intrusive load monitoring in households", **Scientific Data** **2020 7:1**, v. 7, n. 1, p. 1–17, 2 abr. 2020. DOI: 10.1038/s41597-020-0434-6. Disponível em: <https://www.nature.com/articles/s41597-020-0434-6>. Acesso em: 7 dez. 2023.

KOLTER, J. Z., JOHNSON, M. J. "REDD: A Public Data Set for Energy Disaggregation Research", **SustKDD workshop**, v. xxxxx, n. 1, p. 1–6, 2011. Disponível em: <http://www.enmetric.com>. Acesso em: 22 nov. 2023.

LIU, Y., CHEN, M. "A review of nonintrusive load monitoring and its application in commercial building", **4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent Systems, IEEE-CYBER 2014**, p. 623–629, 7 out. 2014. DOI: 10.1109/CYBER.2014.6917536. . Acesso em: 25 fev. 2024.

NOURY, N., BERENGUER, M., TEYSSIER, H., *et al.* "Building an index of activity of inhabitants from their activity on the residential electrical power line", **IEEE**

Transactions on Information Technology in Biomedicine, v. 15, n. 5, p. 758–766, set. 2011. DOI: 10.1109/TITB.2011.2138149. . Acesso em: 1 mar. 2024.

PEREIRA, L., NUNES, N. "Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—A review", **WIRES, Data Mining and Knowledge Discovery**, 2018. DOI: <https://doi.org/10.1002/widm.1265>. .

PUJIĆ, D., JELIĆ, M., TOMAŠEVIĆ, N., *et al.* "Case study from the energy domain", **Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)**, v. 12072 LNCS, p. 165–180, 2020. DOI: 10.1007/978-3-030-53199-7_10/TABLES/1. Disponível em: https://link.springer.com/chapter/10.1007/978-3-030-53199-7_10. Acesso em: 12 dez. 2023.

RASHID, H., SINGH, P., STANKOVIC, V., *et al.* "Can non-intrusive load monitoring be used for identifying an appliance's anomalous behaviour?", **Applied Energy**, v. 238, p. 796–805, 15 mar. 2019. DOI: 10.1016/J.APENERGY.2019.01.061. . Acesso em: 1 mar. 2024.

REHMAN, A. U., LIE, T. T., VALLÉS, B., *et al.* "Comparative Evaluation of Machine Learning Models and Input Feature Space for Non-intrusive Load Monitoring", **Journal of Modern Power Systems and Clean Energy**, v. 9, n. 5, p. 1161–1171, 1 set. 2021. DOI: 10.35833/MPCE.2020.000741. . Acesso em: 28 fev. 2024.

SADEGHIANPOURHAMAMI, N., RUYSSINCK, J., DESCHRIJVER, D., *et al.* "Comprehensive feature selection for appliance classification in NILM", **Energy and Buildings**, v. 151, p. 98–106, 15 set. 2017. DOI: 10.1016/J.ENBUILD.2017.06.042. . Acesso em: 28 fev. 2024.

SCHIRMER, P. A., MPORAS, I. "Non-Intrusive Load Monitoring: A Review", **IEEE Transactions on Smart Grid**, v. 14, n. 1, p. 769–784, 1 jan. 2023. DOI: 10.1109/TSG.2022.3189598. . Acesso em: 19 nov. 2023.

ZOHA, A., GLUHAK, A., IMRAN, M., *et al.* "Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey", **Sensors**, v. 12, n. 12, p. 16838–16866, 6 dez. 2012. DOI: 10.3390/s121216838. .

APÊNDICE A – PROGRAMAÇÃO IMPLEMENTADA NO NILMTK

1º Cenário: TopK 10, 2D

```
In [ ]: #Importa as bibliotecas utilizadas

import numpy as np
import pandas as pd
from os.path import join

from pylab import rcParams
import matplotlib.pyplot as plt
from matplotlib.dates import DateFormatter, DayLocator, HourLocator, MinuteLocator
import matplotlib.dates as mdates
from matplotlib.ticker import MaxNLocator
from datetime import datetime
import pytz
from pytz import timezone

import nilmtk
from nilmtk import DataSet, TimeFrame, MeterGroup, HDFDataStore
from nilmtk.legacy.disaggregate import CombinatorialOptimisation
from nilmtk.utils import print_dict
from nilmtk.metrics import f1_score

import warnings
warnings.filterwarnings("ignore")
```

```
In [ ]: #Definie parâmetros dos gráficos

%matplotlib inline
rcParams['figure.figsize'] = (16, 8)

fonte_titulo = 22
fonte_legenda = 18
fonte_eixo = 18
fonte_letras_eixo = 18
```

```
In [ ]: #Define o arquivo do banco de dados analisado

dataset = DataSet('redd.h5')
```

```
In [ ]: #Define o número do prédio analisado

building_number = 1
elec = dataset.buildings[building_number].elec
mains = dataset.buildings[building_number].elec.mains()
```

Definição das janelas de treinamento e desagregação

```
In [ ]: #Define a janela de dados Ground Truth (A ser desagregados)

inicio_gt = ('2011-04-20')
final_gt = ('2011-04-27')
```

```
In [ ]: #Define a janela de Treinamento do Algoritmo

        inicio_treino = ('2011-04-28')
        final_treino = ('2011-04-30')
```

Definição das cargas que serão desagregadas

```
In [ ]: #Definição das cargas que mais consumiram > Em todo o período disponível < (TopK)

        dataset = DataSet('redd.h5')
        cargas_desag = elec.submeters().select_top_k(k=10)
```

```
In [ ]: #Exibe as cargas que mais consumiram

        print(cargas_desag)
```

```
In [ ]: # Exibe qual o consumo registrado para as cargas que mas consumiram (kWh)

        for meter in cargas_desag.meters:
            energy_consumed = meter.total_energy()
            print(f"{meter}: {energy_consumed} kWh")
```

Inicia o Algoritmo CO

```
In [ ]: #Estabelece a janela de treinamento

        dataset.set_window(start=(inicio_treino),end=(final_treino))
```

```
In [ ]: # Inicia o módulo CombinatorialOptimisation

        co = CombinatorialOptimisation()
```

Fase de Treinamento

```
In [ ]: #Realiza a etapa de treinamento com as cargas escolhidas

        co.train(cargas_desag)
```

Fase de Desagregação

```
In [ ]: #Acessa o banco de dados REDD.h5
        #Estabelece a janela de desagregação

        dataset = DataSet('redd.h5')
        dataset.set_window(start=(inicio_gt),end=(final_gt))
        elec = dataset.buildings[building_number].elec
```

```
In [ ]: #Necessário chamar essa função eventualmente devido bugs no processo de manipulação

        #output.close()
```

```
In [ ]: #Etapa de desagregação, com base nas cargas selecionadas
        #Salva os dados desagregados em outro banco de dados
```

```
disag_filename = ('Disag_TCC_TopK_10_2D.h5')
output = HDFDataStore(disag_filename, 'w')
co.disaggregate(elec.mains(), output)
output.close()
```

```
In [ ]: #Abre o banco de dados resultante da desagregação
```

```
disag = DataSet(disag_filename)
dataset_disag = disag.buildings[1].elec
```

```
In [ ]: #Plot do gráfico comparando os resultados da desagregação da geladeira com os dados
```

```
# Definir a variável para o máximo no eixo y
max_potencia_y = 500 # Altere este valor conforme necessário

instance = 5 # Instancia do aparelho individual a ser plotado
meter_model = dataset_disag[instance]
meter_gt = elec[instance]

# Definir os limites do eixo x em datas de calendário (por exemplo, de 2023-01-01 a
limite_inferior = datetime(2011, 4, 22, 4)
limite_superior = datetime(2011, 4, 22, 22)

# Acessar o fuso horário do dataset
timezone_gt = dataset.metadata['timezone']

# Verificar se ambos os índices estão usando o mesmo fuso horário
if meter_model.power_series_all_data().index.tz != pytz.timezone(timezone_gt):
    # Remover o fuso horário dos dados do modelo
    meter_model.power_series_all_data().index = meter_model.power_series_all_data().

# Plotar o modelo
plt.plot(meter_model.power_series_all_data(), label='Modelo CO')

# Plotar os dados reais
plt.plot(meter_gt.power_series_all_data(), label='Dados Reais')

# Configurar o título
plt.title(f'Resultado da Desagregação: Geladeira - Período: 23/05 a 25/05', fontsize=12)

plt.xlabel('Tempo', fontsize=fonte_eixo)
plt.ylabel('Potência (W)', fontsize=fonte_eixo)

# Definir os limites do eixo x em datas de calendário
plt.xlim(limite_inferior, limite_superior)

# Limitar o eixo y
plt.ylim(0, max_potencia_y)

# Definir o formato desejado para os rótulos de tempo
date_format = DateFormatter('%H:%M')

# Aplicar o formato aos rótulos de tempo no eixo x
plt.gca().xaxis.set_major_formatter(date_format)

# Ajustar o tamanho da fonte dos rótulos dos eixos P e T
plt.xticks(fontsize=fonte_letras_eixo)
plt.yticks(fontsize=fonte_letras_eixo)

# Criar a Legenda apenas para as entradas que têm dados
handles, labels = plt.gca().get_legend_handles_labels()
new_handles = []
```

```

new_labels = []

for handle, label in zip(handles, labels):
    if label in ['Modelo CO', 'Dados Reais']:
        new_handles.append(handle)
        new_labels.append(label)

# Criar a Legenda com as entradas ajustadas
plt.legend(new_handles, new_labels, fontsize=fonte_legenda)

# Mostrar o gráfico
plt.show()

```

F-Score

```

In [ ]: #Plot do gráfico de barras com o resultado do F1-Score

f1 = f1_score(dataset_disag, cargas_desag)
f1.index = dataset_disag.get_labels(f1.index)

# Traduzir os rótulos dos aparelhos
traducao_legendas = {
    'Fridge': 'Geladeira',
    'Air conditioner': 'Ar Condicionado',
    'Sockets': 'Tomadas',
    'Washer dryer': 'Lavadora e Secadora de Roupas',
    'Dish washer': 'Lava-louças',
    'Microwave': 'Micro-ondas',
    'Electric stove': 'Fogão Elétrico',
    'Electric oven': 'Forno Elétrico',
    'Unknown': 'Cargas Desconhecidas',
    'Site meter': 'Energia Agregada',
    'Light': 'Iluminação'
}

f1.index = f1.index.map(traducao_legendas)

# Criar um array de índices para o posicionamento dos rótulos no gráfico
x = np.arange(len(f1))

# Plotar o gráfico de barras
plt.bar(x, f1)
plt.xlabel('Cargas', fontsize=fonte_eixo)
plt.ylabel('F-Score', fontsize=fonte_eixo)
plt.title('F-Score por Carga (K=10) - Treinamento de 28 a 30/04', fontsize=fonte_titulo)

# Adicionar os rótulos dos aparelhos
plt.xticks(x, f1.index, rotation=45, ha="right", fontsize=fonte_letras_eixo)

# Ajustar o tamanho dos ticks do eixo y
plt.yticks(fontsize=fonte_letras_eixo)

# Ajustar o tamanho da Legenda
plt.legend(fontsize=fonte_legenda)

# Exibir o gráfico
plt.show()

```

```

In [ ]: #Exibe os resultados do F1-Score em lista
f1

In [ ]: #Exporta os resultados obtidos para uma planilha usada para gerar o gráfico comparativo
f1.to_csv('cenario1.csv', header=['f_score'])

In [ ]: # Especificar a janela de tempo desejada para o próximo gráfico

limite_inferior = datetime(2011, 4, 22, 16, 0)
limite_superior = datetime(2011, 4, 22, 16, 45)

In [ ]: #Plot do gráfico com o resultado da desagregação de todas as cargas num período definido

# Tabela de tradução dos rótulos
traducao_legendas = {
    'Fridge': 'Geladeira',
    'Air conditioner': 'Ar Condicionado',
    'Sockets': 'Tomadas',
    'Washer dryer': 'Lavadora e Secadora de Roupas',
    'Dish washer': 'Lava-louças',
    'Microwave': 'Micro-ondas',
    'Electric stove': 'Fogão Elétrico',
    'Electric oven': 'Forno Elétrico',
    'Unknown': 'Cargas Desconhecidas',
    'Site meter': 'Energia Agregada',
    'Light': 'Iluminação'
}

# Gráfico com dados do modelo
plt.figure()
ax1 = plt.gca()
for instancia in instancias_desejadas:
    try:
        appliance_name = instances_labels[instancia]
        appliance_name_pt = traducao_legendas.get(appliance_name, appliance_name)
        appliance_series = dataset_disag[instancia].power_series_all_data()
        appliance_series = appliance_series.loc[limite_inferior:limite_superior]
        ax1.plot(appliance_series.index, appliance_series.values, label=appliance_name_pt)
    except KeyError:
        print(f"Instância {instancia} não encontrada no conjunto de dados desagregados")

ax1.xaxis.set_major_locator(MinuteLocator(interval=30))
ax1.xaxis.set_major_formatter(DateFormatter('%H:%M:%S', tz=pytz.timezone('America/Moscow')))

plt.gcf().autofmt_xdate()

# Ajustar o tamanho dos ticks do eixo y
plt.yticks(fontsize=fonte_letras_eixo)
# Ajustar o tamanho dos ticks do eixo x
plt.xticks(fontsize=fonte_letras_eixo)
plt.xlabel('Tempo', fontsize=fonte_eixo)
plt.ylabel('Potência (W)', fontsize=fonte_eixo)
plt.title('Detalhe da potência desagregada no momento \n de acionamento do compressor')

legenda_ax1 = ax1.legend(fontsize=fonte_legenda, bbox_to_anchor=(0.5, -0.2), loc="upper right")
if legenda_ax1:
    posicao_legenda = legenda_ax1.get_bbox_to_anchor().inverse_transformed(ax1.transData)
    plt.savefig('grafico_modelo.png', bbox_inches='tight', pad_inches=0.5)
else:
    plt.savefig('grafico_modelo.png', bbox_inches='tight', pad_inches=0.5)
plt.show()

```



```
In [ ]: #Função Help, disponível para todos os módulos do NILMTK
        help(CombinatorialOptimisation)
```

Gráfico Comparativo de Todos os Cenários

```
In [ ]: from itertools import cycle
        import seaborn as sns
```

```
In [ ]: # Ler cada CSV e criar um DataFrame
        dfs = [pd.read_csv(caminho, index_col=0, header=None, names=['f_score']) for caminho in caminhos]

        # Traduzir os índices conforme a Lista de tradução
        dfs = [df.rename(index=traducao_legendas) for df in dfs]
```

```
In [ ]: # Adicionar uma coluna de cenário em cada DataFrame
        cenarios = ['2 Dias de Treino (10 cargas)', '2 Dias de Treino (13 cargas)',
                    'Terceiro Cenário (2 dias)', '1 Dia de Treino (10 cargas)',
                    '1 Dia de Treino (13 cargas)']
```

```
In [ ]: # Traduzir os índices conforme a Lista de tradução
        dfs = [df.rename(index=traducao_legendas) for df in dfs]
```

```
In [ ]: # Adicionar uma coluna de cenário em cada DataFrame
        cenarios = ['2 Dias de Treino (10 cargas)', '2 Dias de Treino (13 cargas)',
                    'Terceiro Cenário (2 dias)', '1 Dia de Treino (10 cargas)',
                    '1 Dia de Treino (13 cargas)']

        dfs = [df.assign(Cenário=cenario) for df, scenario in zip(dfs, cenarios)]
```

```
In [ ]: # Concatenar os DataFrames em um único DataFrame
        data = pd.concat(dfs).reset_index()
```

```
In [ ]: # Definir parâmetros dos gráficos
        %matplotlib inline
        rcParams['figure.figsize'] = (16, 8)

        fonte_titulo = 22
        fonte_legenda = 18
        fonte_eixo = 18
        fonte_letras_eixo = 18

        # Configurar o estilo do Seaborn
        sns.set(style="whitegrid")
        sns.set_context("paper", rc={"lines.linewidth": 0.7}) # Ajuste da Largura da Linha

        # Carregar os dados de cada cenário
        cenarios = ['cenario1', 'cenario2', 'cenario3', 'cenario4', 'cenario5']
        dfs = [pd.read_csv(f'{cenario}.csv') for scenario in cenarios]

        # Traduzir os rótulos das cargas
        traducao_legendas = {
            'Geladeira': 'Geladeira',
            'Lavadora e Secadora de Roupas': 'Lavadora/Secadora',
            'Iluminação': 'Iluminação',
            'Tomadas': 'Tomadas',
```

```

    'Lava-louças': 'Lava-louças',
    'Micro-ondas': 'Micro-ondas',
    'Forno Elétrico': 'Forno Elétrico',
    'Cargas Desconhecidas': 'Cargas Desconhecidas',
    'Fogão Elétrico': 'Fogão Elétrico',
    'Aquecedor de Ambientes': 'Aquecedor Ambientes'
}

nomes_cenarios = {
    'cenario1': 'Treinamento 2 Dias - TopK 10',
    'cenario2': 'Treinamento 2 Dias - Top K 13',
    'cenario3': 'Teinamento 2 Dias - Todas as Cargas',
    'cenario4': 'Treinamento 1 Dia - Top K 10',
    'cenario5': 'Treinamento 1 Dia - Top K 13'
}

# Criar um DataFrame com os dados
data = pd.DataFrame(columns=['Cargas', 'F-Score', 'Cenário'])

# Adicionar os dados de cada cenário ao DataFrame
for i, cenario in enumerate(cenarios):
    df = dfs[i]
    data = pd.concat([data, pd.DataFrame({
        'Cargas': [traducao_legendas[carga] for carga in df['Unnamed: 0']],
        'F-Score': df['f_score'],
        'Cenário': nomes_cenarios[cenario]
    })])

# Definir paletas personalizadas para azuis e verdes
paleta_azuis = sns.color_palette("Blues", 4)
paleta_verdes = sns.color_palette("Greens", 4)

# Mapear as cores para cada cenário
mapeamento_cores = {
    nomes_cenarios['cenario1']: paleta_azuis[2],
    nomes_cenarios['cenario2']: paleta_azuis[3],
    nomes_cenarios['cenario3']: paleta_azuis[0],
    nomes_cenarios['cenario4']: paleta_verdes[2],
    nomes_cenarios['cenario5']: paleta_verdes[3]
}

# Criar o gráfico com cores distintas para os cenários 1, 2 e 3, e para os cenários
plt.figure(figsize=(16, 8))

# Adicionar o parâmetro capsiz para incluir barras de erro representando os valores
ax = sns.barplot(x='Cargas', y='F-Score', hue='Cenário', data=data, palette=mapeame

# Rotacionar as etiquetas no eixo x
plt.xticks(rotation=45, ha='right')

# Configurar o título e os rótulos dos eixos
plt.title('F-Score por Cenário e Cargas', fontsize=fonte_titulo)
plt.xlabel('Cargas', fontsize=fonte_eixo)
plt.ylabel('F-Score', fontsize=fonte_eixo)

# Adicionar a legenda no canto superior direito
plt.legend(title='Cenário', loc='upper right', fontsize=fonte_legenda)

# Ajustar o tamanho das letras no eixo x e y
plt.xticks(fontsize=fonte_letras_eixo)
plt.yticks(fontsize=fonte_letras_eixo)

# Mostrar o gráfico
plt.show()

```

```
In [ ]: cenarios = ['cenario1', 'cenario2', 'cenario3', 'cenario4', 'cenario5']

for cenario in cenarios:
    print(f'Cenário: {cenario}')
    df = pd.read_csv(f'{cenario}.csv')
    print(df)
    print('\n' + '-' * 50 + '\n')
```

```
In [ ]:
```