

Bases de dados para Monitoramento Não-intrusivo da Carga: uma revisão

Non-Intrusive Load Monitoring Dataset: a review

Leandro de Freitas Velozo

Programa de Pós-Graduação em Gestão de Redes de
Telecomunicações
Pontifícia Universidade Católica de Campinas
Campinas, Brasil
leandro.fv@puc Campinas.edu.br

Lia Toledo Moreira Mota

Programa de Pós-Graduação em Gestão de Redes de
Telecomunicações
Pontifícia Universidade Católica de Campinas
Campinas, Brasil
lia.mota@puc-campinas.edu.br

Abstract— Currently there are a lot of efforts to improve Electrical Energy Efficiency and saving natural resource. In order to support this effort, there are many research related to Non-Intrusive Load Monitoring (NILM). With this technique, it is possible to analyze electrical power consumption of individual loads, also called disaggregation of loads, using only a meter typically installed in the consumer energy input. NILM used in residence can bring some advantages for the customers in order to identify the energy consumption behavior of individual equipment. The NILM have been used with intelligent algorithms which need to be training, using datasets to validate their performance. Worldwide, there are about forty-two Datasets but in Brazil, there is only one. Considering the importance of NILM Datasets for load desegregation, this presents a critical review of NILM Datasets and as results the challenge for future research.

Keywords—Non-Intrusive Load Monitoring; Dataset

I. INTRODUÇÃO

Com o aumento da população mundial e o desenvolvimento socioeconômico, tem-se um aumento do uso dos recursos naturais, que em geral são finitos. Mediante esse cenário, a busca por eficiência energética com objetivo de reduzir o consumo de recursos naturais é uma área de pesquisa em franco desenvolvimento, liderada, principalmente, na área de sistemas de energia elétrica, por pesquisas relacionadas a *Smart Grids*. No contexto das Redes Elétricas Inteligentes, se inserem, também, as pesquisas relacionadas a desagregação de cargas residenciais. A desagregação de cargas residenciais tem como objetivo principal de informar ao usuário o perfil de consumo individualmente dos equipamentos elétricos e eletrônicos, contribuindo para um uso consciente e com maior eficiência. De acordo com o Plano Decenal de Expansão de Energia 2031, publicado pela Empresa de Pesquisas Energéticas – EPE [1], o consumo total de eletricidade no Brasil deve ter um crescimento médio de 3,5% anualmente entre 2021 e 2031, para um desempenho esperado do produto interno bruto de 2,9% anual no período. Nesse estudo, estima-se que o consumo médio das residências deve passar de 165 kWh/mês em 2021 para 196 kWh/mês em 2031. Corroborando o estudo, foi observado que devido aos avanços tecnológicos, houve mudança significativa no perfil de usuários residenciais ao longo dos anos e que, atualmente, tem-se um número significativo de dispositivos e

equipamentos eletrônicos nas residências, o que ressalta a importância de estudos na desagregação de cargas residenciais. O conceito de desagregação de cargas é explorado na literatura a partir de três diferentes metodologias, sendo elas o monitoramento de cargas intrusivo, em geral com uso de *smart plugs* para monitoramento individual de equipamentos residenciais; o monitoramento de cargas semi-intrusivo, que é denominado assim devido aos pontos de medição serem em alguns circuitos específicos; e o monitoramento de cargas não-intrusivo, que é feito através de um único ponto de medição, usualmente sendo esse ponto na entrada do quadro de distribuição.

O monitoramento de cargas residenciais não-intrusivo, conhecido como *Non-Intrusive Load Monitoring (NILM)* tem sido objetivo de várias pesquisas devido à sua contribuição em eficiência energética. Para o consumidor residencial, a possibilidade de conhecer o perfil de consumo individual de cada equipamento ou eletrodoméstico em sua residência, fornece a possibilidade de fazer um gerenciamento de uso, e inclusive ajuda na tomada de decisão de substituição do equipamento ou eletrodoméstico por outro com melhor eficiência energética.

A metodologia NILM foi concebida, inicialmente, baseada na teoria de conservação de energia. Com os avanços computacionais e novos algoritmos, os sistemas NILM passaram a usar ferramentas computacionais como *Machine Learning* para fazer a desagregação de cargas residenciais a partir da medição em um único ponto.

Para um desempenho adequado do sistema NILM, os algoritmos existentes necessitam de uma base de dados, também conhecida na literatura como *Dataset* para fazer o aprendizado e, posteriormente, fazer o reconhecimento de padrões de carga. As bases de dados contêm grandezas como tensão, corrente e potência de equipamentos residenciais. As bases de dados são, também, de extrema importância para pesquisadores que estão desenvolvendo novos algoritmos de desagregação de cargas.

Nesse contexto, esse trabalho tem como objetivo fazer uma revisão sistemática das Bases de Dados para NILM disponíveis e identificar os principais desafios para a implementação de uma base de dados nacional, acessível aos pesquisadores da área. O trabalho está dividido em cinco seções. A Seção II apresenta um breve estudo sobre sistema de monitoramento de cargas não-intrusivo. A Seção III apresenta uma revisão biblio-

gráfica sobre as principais bases de dados públicas disponíveis mundialmente. Finalmente, a Seção IV aborda uma discussão sobre os requisitos para uma nova base de dados nacional.

II. ESTADO DA ARTE EM NILM

No início dos anos 80, no Instituto de Tecnologia de Massachusetts - MIT, um trabalho desenvolvido por George W. Hart [2] com o apoio de Ed Kern e Fred Schweppe, sendo financiado pelo *Electric Power Research Institute*, foi um dos pioneiros nessa área. Esse trabalho tratava da medição não-intrusiva de utensílios domésticos (*Non-intrusive Appliance Load Monitoring NIALM*). Basicamente, esse trabalho atesta que, a partir da medição de potência ativa e reativa em momentos de ligar e desligar os equipamentos, seja possível fazer a distinção entre eletrodomésticos com cargas semelhantes, através das diferentes impedâncias complexas. O sistema NILM está dividido nas etapas de aquisição de dados, detecção de eventos, extração de dados e identificação de carga, conforme descrito a seguir.

A. Aquisição de Dados

A aquisição de dados é o estágio onde é feita a medição de tensão e corrente em geral e esses sinais são digitalizados para o processamento e cálculos de grandezas elétricas.

B. Detecção de Eventos

A etapa de detecção de evento tem como objetivo principal identificar o momento em que a carga elétrica foi ligada ou desligada. Esses pontos são de extrema importância para a identificação e posteriormente a desagregação de cargas.

C. Extração de Dados

Este estágio vai fazer uso dos dados coletados, como tensão, corrente e formas de onda para calcular e identificar potência ativa, reativa e harmônicos de correntes.

D. Identificação de Cargas

Através de algoritmos de reconhecimento de padrões, é possível com os dados coletados e extraídos, identificar as cargas em momentos em que foram ligadas e desligadas.

A figura 1 apresenta um exemplo de desagregação de cargas baseada na detecção de eventos. É apresentado a potência ativa consumida num período de quarenta minutos. No exemplo, tem-se quatro diferentes níveis de potência bem definidos. No monitoramento de cargas, é possível analisar os dados de medição em dois estados, o estado de regime permanente (*steady-state*) que são cargas que têm estados de operação muito bem definidos em condições de ligado e desligado, como apresentado na figura 1, e o estado transiente ou (*transient state*), em que são classificados os equipamentos que têm estados de operação variáveis como, por exemplo, máquinas de lavar roupa que têm ciclos de operações.

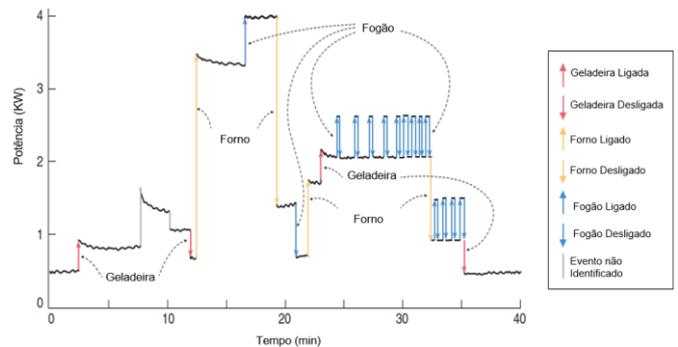


Fig. 1. Exemplo desagregação de cargas baseada em eventos [1] [2]

Para equipamentos operando em regime permanente, as grandezas corrente, potência ativa e reativa e admitância são suficientes para o NILM, porém para os equipamentos em estado transiente, são necessários mais dados para identificação da “assinatura de carga”, sendo eles análise espectral, resposta em frequência e transformada Wavelet, conforme apresentado no diagrama da figura 2.

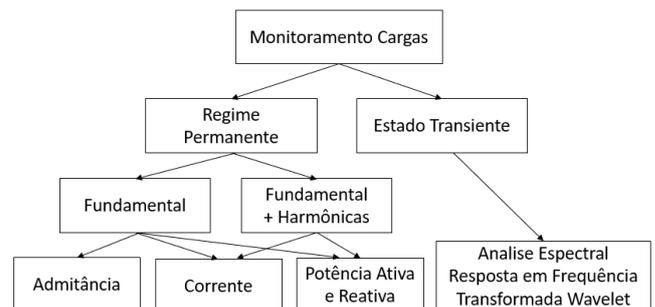


Fig. 2. Classificação de métodos NILM [3]

As pesquisas sobre desagregação de cargas residenciais ficaram durante um período de quase vinte anos desde o trabalho de Hart [2] com poucas publicações nessa área. Entretanto, com a evolução dos dispositivos de aquisição de dados e, principalmente, com técnicas de inteligência artificial, as pesquisas nessa área voltaram a ganhar destaque e, com isso, os sistemas de desagregação de cargas com *Machine Learning* e *Deep Learning Algorithms* são campos bem atrativos de pesquisa atualmente. Para que o sistema NILM atue de forma adequada, uma base de dados composta por grandezas relacionadas às cargas é extremamente importante. Na figura 3, tem-se um diagrama de um sistema NILM com uso de algoritmos inteligentes. O uso do sistema NILM com algoritmos inteligentes está separado em duas etapas, a etapa de treinamento e a etapa de monitoramento. Na etapa de treinamento, após a aquisição e extração de eventos, o algoritmo vai ser treinado a identificar e classificar as cargas, usando uma base de dados, que é de extrema importância para a precisão do sistema. Concluída essa etapa, o sistema vai estar apto a funcionar na etapa de monitoramento, onde o sistema vai fazer a desagregação de cargas continuamente.

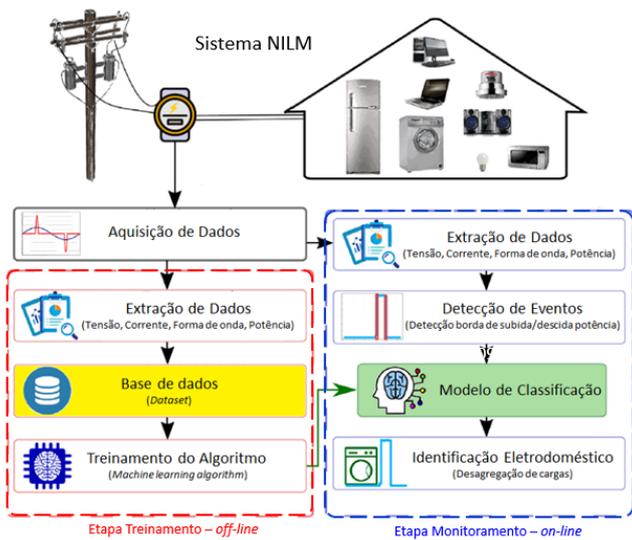


Fig. 3. Diagrama NILM com uso de algoritmos inteligentes [4]

Conforme apresentado na figura 3, para o funcionamento do sistema, é requerida uma base de dados ou *dataset*, que deve conter curvas características de equipamentos e eletrodomésticos, que geralmente são curvas de tensão e corrente adquiridas em alta frequência e com alta resolução para o algoritmo a ser treinado com esses padrões. Essas curvas são também chamadas de “assinatura de carga”. Após essa etapa, o sistema vai estar apto a reconhecer e desagregar equipamentos e eletrodomésticos, também, em situações em que tem-se associação de cargas ligadas.

III. BASE DE DADOS – REVISÃO BIBLIOGRÁFICA

As bases de dados são importantes para diversas áreas de pesquisa. Elas permitem os pesquisadores possam reproduzir cenários, fazer testes e experimentos para provas de conceitos. As bases de dados para NILM são ferramentas importantes no desenvolvimento, aprendizado e validação de algoritmos e, nos últimos anos, devido a sua importância e necessidade, têm sido objetivo de estudos de vários pesquisadores [7-17]. Um recente trabalho publicado com uma revisão do estado da arte de base de dados para NILM de Hafiz et al. [5] identificou quarenta e duas bases de dados públicas.

A figura 4 apresenta o diagrama de categorias em que são classificadas as bases de dados. Primeiramente, estão classificadas em relação à frequência, sendo considerada baixa frequência, aquelas que têm dados disponíveis com taxa de amostragem em Hz, e alta frequência para dados na ordem de kHz. Uma segunda classificação é referente ao tipo de carga: as cargas naturais são cargas que têm os dados coletados em ambiente real de uso e as cargas sintéticas, em geral, são dados de cargas coletadas em ambiente controlado, por exemplo, em laboratórios com bancada de teste. E também tem-se as cargas simuladas, que estão pouco presentes nas bases de dados disponíveis publicamente. Para essas cargas simuladas, não existe um consenso quanto à sua utilização. Com relação às

cargas naturais, pode-se encontrar bases de dados com dados individuais das cargas e ainda dados de cargas associadas.

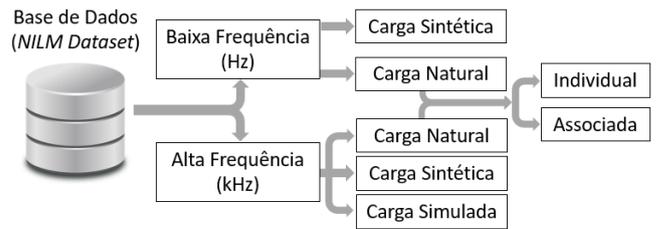


Fig. 4. Classificação de base de dados para sistema NILM

A tabela 1 apresenta as bases de dados públicas disponíveis com a data de criação, país de origem, frequência, ambiente da coleta de dados e números de ambientes.

TABELA 1 – BASES DE DADOS

| Ano | Base de Dados | País | Frequência | | Ambiente | Nº de Ambiente(s) |
|------|---------------|--------------------|------------|------|-------------|-------------------|
| | | | Baixa | Alta | | |
| 2011 | REDD | Estados Unidos | ✓ | ✓ | Residencial | 6 |
| 2012 | Smart | Estados Unidos | ✓ | | Residencial | 7 |
| 2012 | BLUED | Estados Unidos | | ✓ | Residencial | 1 |
| 2012 | HES | Reino Unido | ✓ | | Residencial | 251 |
| 2012 | Tracebase | Alemanha | ✓ | | Residencial | 15 |
| 2013 | BERDS | Estados Unidos | ✓ | | Comercial | 1 |
| 2013 | Dataport | Estados Unidos | ✓ | | Res & Com* | 722 |
| 2013 | AMPds | Canada | ✓ | | Residencial | 1 |
| 2013 | iAWE | India | ✓ | | Residencial | 1 |
| 2013 | IHEPCDS | França | ✓ | | Residencial | 1 |
| 2013 | ACS-Fx | Suíça | ✓ | | Residencial | - |
| 2014 | UK-DALE | Reino Unido | ✓ | | Residencial | 5 |
| 2014 | ECO | Suíça | ✓ | | Residencial | 6 |
| 2014 | GREEND | Austria / Italia | ✓ | | Residencial | 9 |
| 2014 | PLAID I | Estados Unidos | | ✓ | Residencial | 56 |
| 2014 | RBSA | Estados Unidos | ✓ | | Residencial | 101 |
| 2014 | SustData | Portugal | | ✓ | Residencial | 1 |
| 2014 | COMBED | India | ✓ | | Laboratório | 1 |
| 2015 | DRED | Holanda | ✓ | | Residencial | 1 |
| 2015 | HFED | India | | ✓ | Residencial | 1 |
| 2015 | REFIT | Reino Unido | ✓ | | Residencial | 5 |
| 2016 | WHITED | Alemanha / Austria | ✓ | ✓ | Residencial | - |
| 2016 | COOLL | França | | ✓ | Residencial | - |
| 2016 | SustDataED | Portugal | | ✓ | Residencial | 1 |
| 2016 | OPLD | Singapura | ✓ | | Comercial | - |
| 2016 | AMPds2 | Canada | ✓ | | Residencial | 1 |
| 2016 | SmartSim | Estados Unidos | ✓ | | Residencial | 1 |
| 2017 | AMBEL | Alemanha | ✓ | | Residencial | - |
| 2017 | PLAID II | Estados Unidos | | ✓ | Residencial | 9 |
| 2017 | EEUD | Canada | ✓ | | Residencial | 23 |
| 2017 | ESHL | Alemanha | ✓ | | Laboratório | 1 |
| 2018 | RAE | Canada | ✓ | | Residencial | 1 |
| 2018 | SHED | França | ✓ | | Comercial | 8 |
| 2018 | BLOND | Alemanha | | ✓ | Residencial | 1 |
| 2018 | HELD1 | Alemanha | | ✓ | Laboratório | 1 |
| 2019 | ENERTALK | Coreia do Sul | ✓ | | Residencial | 22 |
| 2019 | I-BLEND | India | ✓ | | Res & Com | 7 |
| 2019 | HUE | Canada | ✓ | | Residencial | 22 |
| 2020 | SynD | Austria | ✓ | | Residencial | 2 |
| 2020 | IDEAL | Reino Unido | ✓ | | Residencial | 255 |
| 2020 | CU-BEMS | Tailândia | ✓ | | Comercial | 1 |
| 2020 | PLAID III | Estados Unidos | | ✓ | Residencial | 1 |
| 2021 | LIT-Dataset | Brasil | | ✓ | Residencial | 1 |

*Carga Residencial e Comercial

Além das classificações mencionadas, as bases de dados diferem entre si por outras características como quantidade de ambientes coletados, tempo de aquisição, dados contínuos e não contínuos e as informações coletadas. Essas informações são, em geral, tensão, corrente e potência ativa. Porém, existem bases de dados que disponibilizam mais informações como potência aparente e reativa, fator de potência e admitância.

É possível observar que tem-se um número bem menor de base de dados em alta frequência e notar que as bases de dados estão concentradas nos Estados Unidos e Europa. Na América Latina, em especial no Brasil, tem-se somente uma base de dados.

Outro ponto a ser considerado nas bases de dados existentes está relacionado à forma e formato em que os dados estão disponíveis. Em alguns casos, o acesso e manipulação desses dados não são amigáveis e requerem o uso de scripts para acesso e manipulação de dados.

IV. CONCLUSÕES

Existem vários desafios em relações as bases de dados existentes para sistema NILM, sendo os principais a falta de variedade de dados, uma padronização entre as bases de dados afim de compartilhar dados entre elas, métricas bem definidas de parâmetros para validação de bases de dados.

Para concluir esse estudo, baseado nos desafios identificados na revisão da literatura, propõe-se uma base de dados com cargas naturais residenciais, com equipamentos e eletrodomésticos comercializados no Brasil, em alta frequência de forma individual e associada. A tabela 2 apresenta a proposta de equipamentos para essa base de dados que atualmente está em desenvolvimento.

TABELA 2 – EQUIPAMENTOS DA BASE DE DADOS PROPOSTA

| Item | NOME | MODELO | FABRICANTE | POTÊNCIA |
|------|---|------------------|----------------|----------|
| 1 | AirFryer Viva | R19217 | Philips Walita | 1425W |
| 2 | Aquecedor a óleo Portátil | DeLonghi | EW6507 | 15000W |
| 3 | Ar Condicionado Split Dual Inverter 12000 | S4NQ12JA3WC | LG | 1085W |
| 4 | Ar Condicionado Split Dual Inverter 12000 | 42MBCB09M5 | Springer Midea | 1060W |
| 5 | Aspirador | Sonic | Electrolux | 1400W |
| 6 | Bebedouro com compressor | BC21B | Eletrolux | 130W |
| 7 | Boiler Convencional Alta Pressão | RTP500 LTS 40MCA | Industek | 2500W |
| 8 | Bomba de pressurização e circulação | GP-300P | Inova | 300W |
| 9 | Cervejeira Beer Maxx 250 | VN25TE | Metafrio | 210W |
| 10 | Chuveiro | Acqua duo | Lorenzetti | 7800W |
| 11 | Cooktop de Indução | IC60 | Eletrolux | 7400W |
| 12 | Desktop PC Pavilion Slimline | S5000 | HP | 300W |
| 13 | Ferro elétrico | Easyline SIE60 | Electrolux | 1200W |
| 14 | Forno de embutir Elétrico 80L | OE8DX | Eletrolux | 3000W |
| 15 | Laptop | Z book i7 | HP | 150W |
| 16 | Lava e seca 11Kg | WD11M44 | Samsung | 1100W |
| 17 | Lavadora de alta pressão | Wap | FW004194 | 1750W |
| 18 | Máquina de Café Dolce Gusto Mini Me | DMM6 | Arno | 1460W |
| 19 | Microondas Embutir 28L | MB38T | Eletrolux | 1450W |
| 20 | Motobomba centrífuga residencial | MB71E0005AS | Mundial SA | 850W |
| 21 | Refrigerador Inverse 443 litros | BRE57AKANA | Brastemp | 140W |
| 22 | Secador de cabelo | Taiff Style | Taiff | 2000W |
| 23 | Smart TV 4K NanoCell IPS 65" Ultra HD | 65NANO81 | LG | 200W |
| 24 | Soprador | Maxx | Kiklon | 1400W |
| 25 | Veneziana Elétrica | - | VN21 | 300W |

- [1] Empresa de Pesquisa Energética. Plano Decenal de Expansão de Energia 2031. Disponível em < https://www.epe.gov.br/sites-pt/publicacoes-dadosbertos/publicacoes/Documents/PDE%202031_ReviseoPosCP_rvFinal_v2.pdf>. Acesso em 11 de setembro de 2022.
- [2] L. Pereira, N. Nunes, Performance evaluation in non-intrusive load monitoring: Datasets, metrics, and tools—a review, *Wiley Interdiscip. Rev.:Data Min. Knowl. Discov.* 8 (2018) e1265.
- [3] H. Najmeddine et al., "State of art on load monitoring methods," 2008 IEEE 2nd International Power and Energy Conference, 2008, pp. 1256-1258, doi: 10.1109/PECON.2008.4762669.
- [4] Wesley A. Souza, Augusto M.S. Alonso, Thais B. Bosco, Fernando D. Garcia, Flavio A.S. Gonçalves, Fernando P. Marafão, Selection of features from power theories to compose NILM datasets, *Advanced Engineering Informatics*, Volume 52, 2022, 101556, ISSN 1474-0346, <https://doi.org/10.1016/j.aei.2022.101556>.
- [5] Hafiz Khurram Iqbal, Farhan Hassan Malik, Aoun Muhammad, Muhammad Ali Qureshi, Muhammad Nawaz Abbasi, Abdul Rehman Chishti, A critical review of state-of-the-art non-intrusive load monitoring datasets, *Electric Power Systems Research*, Volume 192, 2021, 106921, ISSN 0378-7796, <https://doi.org/10.1016/j.epsr.2020.106921>.
- [6] Renaux, D.P.B.; Pottker, F.; Ancelmo, H.C.; Lazzaretti, A.E.; Lima, C.R.E.; Linhares, R.R.; Oroski, E.; Nolasco, L.d.S.; Lima, L.T.; Mulinari, B.M.; Silva, J.R.L.d.; Omori, J.S.; Santos, R.B.d. A Dataset for Non-Intrusive Load Monitoring: Design and Implementation. *Energies* 2020, 13, 5371. <https://doi.org/10.3390/en13205371>
- [7] Souza, W.A.; Garcia, F.D.; Marafão, F.P.; da Silva, L.C.P.; Simões, M.G. Load Disaggregation Using Microscopic Power Features and Pattern Recognition. *Energies* 2019, 12, 2641. <https://doi.org/10.3390/en12142641>
- [8] N. Sadeghianpourhamami, J. Ruysinck, D. Deschrijver, T. Dhaene, C. Develder, Comprehensive feature selection for appliance classification in NILM, *Energy and Buildings*, Volume 151, 2017, Pages 98-106, ISSN 0378-7788, <https://doi.org/10.1016/j.enbuild.2017.06.042>.
- [9] Nur Farahin Esa, Md Pauzi Abdullah, Mohammad Yusri Hassan, RETRACTED: A review disaggregation method in Non-intrusive Appliance Load Monitoring, *Renewable and Sustainable Energy Reviews*, Volume 66, 2016, ISSN 1364-0321. <https://doi.org/10.1016/j.rser.2016.07.009>.
- [10] Y. F. Wong, Y. Ahmet Şekercioğlu, T. Drummond and V. S. Wong, "Recent approaches to non-intrusive load monitoring techniques in residential settings," *2013 IEEE Computational Intelligence Applications in Smart Grid (CIASG)*, 2013, pp. 73-79, doi: 10.1109/CIASG.2013.6611501.
- [11] I. Abubakar, S. N. Khalid, M. W. Mustafa, H. Shareef and M. Mustapha, "An overview of Non-intrusive load monitoring methodologies," *2015 IEEE Conference on Energy Conversion (CENCON)*, 2015, pp. 54-59, doi: 10.1109/CENCON.2015.7409513.
- [12] Le, T.-T.-H.; Kim, H. Non-Intrusive Load Monitoring Based on Novel Transient Signal in Household Appliances with Low Sampling Rate. *Energies* 2018, 11, 3409. <https://doi.org/10.3390/en11123409>
- [13] E.J. Aladesanmi, K.A. Folly, Overview of non-intrusive load monitoring and identification techniques, *IFAC-PapersOnLine*, Volume 48, Issue 30, 2015, ISSN 2405-8963, <https://doi.org/10.1016/j.ifacol.2015.12.414>.
- [14] Klemenjak, C. On performance evaluation and machine learning approaches in non-intrusive load monitoring. *Energy Inform 1 (Suppl 1)*, 36 (2018). <https://doi.org/10.1186/s42162-018-0051-1>
- [15] Bao, K., Ibrahimov, K., Wagner, M. et al. Enhancing neural non-intrusive load monitoring with generative adversarial networks. *Energy Inform 1 (Suppl 1)*, 18 (2018). <https://doi.org/10.1186/s42162-018-0038-y>
- [16] Rottondi, C., Derboni, M., Piga, D. et al. An optimisation-based energy disaggregation algorithm for low frequency smart meter data. *Energy Inform 2 (Suppl 1)*, 13 (2019). <https://doi.org/10.1186/s42162-019-0089-8>
- [17] R. Gopinath, Mukesh Kumar, C. Prakash Chandra Joshua, Kota Srinivas, Energy management using non-intrusive load monitoring techniques – State-of-the-art and future research directions, *Sustainable Cities and Society*, Volume 62, 2020, ISSN 2210-6707, <https://doi.org/10.1016/j.scs.2020.102411>.