On Energy Theft Attack Detection in Smart Grids Using Machine Learning

Ahlam Saud Althobaiti, BSc (Hons), MRes
School of Computing and Communications
Lancaster University

A thesis submitted for the degree of
Doctor of Philosophy

November, 2023

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and belief, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of $80,000$ words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: 36781

Ahlam Saud Althobaiti

On Energy Theft Attack Detection in Smart Grids Using Machine Learning

Ahlam Saud Althobaiti, BSc (Hons), MRes.
School of Computing and Communications, Lancaster University
A thesis submitted for the degree of *Doctor of Philosophy*. November, 2023

Abstract

The convergence of legacy power system components with advanced information and communication facilities has led to the emergence of smart grids. Smart grids are envisioned to be the next generation of innovative power systems, guaranteeing resilience, reliability and sustainability, and facilitating energy production, distribution and management. Nonetheless, the development of such systems entails challenges covering a broad spectrum, ranging from operational management to data-driven power accounting and network security. Given the highly distributed properties of the modern grid, energy theft attacks can now be observed at various transmission and distribution levels. Apart from the financial gains for malicious actors, energy theft can also affect critical grid processes and have a direct impact on the grid's overall resilience and safety. Conventional energy theft detection approaches rely on physically inspections, which are time-consuming, inaccurate, costly and require substantial human labour. By virtue of the smart grid paradigm, these inspections are now conducted more efficiently using modern data-driven and machine learning-based detection approaches. Therefore, the major focus of this thesis is on designing a data-driven energy theft detection framework, taking advantage of the unique characteristics of modern smart grids. In particular, this thesis investigates and surveys the advances in energy theft strategies, as well as detection methods, from different perspectives on the smart grid, revolving around energy data manipulation of all three functions of demand, supply and generation. In addition, this thesis proposes a supervisory control and data acquisition (SCADA)-agnostic power modelling scheme for distributed renewable energy sources (DRES). Through this study, it is demonstrated that a viable and exogenous profiling solution achieving similar accuracy to SCADA-based schemes but under much lower computational time is required to produce a reliable regression model for DRES generation energy. Building on this work on SCADA-agnostic DRES power modelling, this thesis also describes a predictive energy theft detection approach for DRES. Evidently, the proposed approach yields a high DRES-based energy theft detection accuracy rate of over 95%, with low computational time required to produce DRES theft classification. Thus, it reasonably addresses the highly demanding requirements of low-cost and accurate real-time energy theft detection in modern power grids. Finally, this thesis introduces

a self-learning theft detection system capable of distinguishing the properties of power consumption and generation theft with possible misconfigurations caused by non-malicious intent. The proposed approach is adaptive through a self-learning operation that is continuously updated as new measurements become available. The results obtained indicate that this scheme can achieve over 90% accuracy in identifying theft with optimal over-streamed data measurements. Thus, it offers low computational time being required to classify consumption and generation meters, and its properties can be exploited for next-generation cross-batch energy theft detection.

Publications

The following publication have been generated while developing this thesis, and to an extent has guided the thesis into what it has become:

Ahlam Althobaiti, Anish Jindal, and Angelos K Marnerides (2020). "Scada-agnostic power modelling for distributed renewable energy sources". In: *2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. IEEE, pp. 379–384

Ahlam Althobaiti, Anish Jindal, and Angelos K. Marnerides (2021). "Data-Driven Energy Theft Detection in Modern Power Grids". In: *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. e-Energy '21. New York, NY, USA: Association for Computing Machinery, pp. 39–48

Ahlam Althobaiti, Anish Jindal, Angelos K Marnerides, and Utz Roedig (2021). "Energy Theft in Smart Grids: A Survey on Data-Driven Attack Strategies and Detection Methods". In: *IEEE Access* 9, pp. 159291–159312

Ahlam Althobaiti, Charalampos Rotsos, and Angelos K Marnerides (2023). "Adaptive Energy Theft Detection in Smart Grids Using Self-Learning With Dual Neural Network". In: *IEEE Transactions on Industrial Informatics*

Acknowledgements

First and foremost, I would like to express my gratitude to Allah (Almighty God) for everything. In the following lines, I wish to convey my appreciation to everyone who has played a pivotal role in making my time as a PhD candidate at Lancaster University both valuable and enjoyable.

My sincere appreciation and gratitude go to Prof. Angelos Marnerides, my primary supervisor. He gave me patient guidance, encouragement and invaluable advice, all while allowing me the freedom to work in my own way. Even after he left Lancaster University to work at Glasgow University, he continued to offer his unwavering support and guidance. I would also like to extend my thanks to Dr Charalampos Rotsos, my secondary supervisor, who invested much time and effort in helping me develop my work and graciously responded to my numerous questions and concerns. They have both been a genuine source of inspiration, and I will forever be indebted to them.

I also want to extend my love and gratitude to my mother, Mrs Rabiha, and my father, Mr Saud, to whom I dedicate my PhD thesis and all my endeavours throughout my lifetime. I wish to convey my deepest appreciation for their unwavering support, love, faith in me and prayers. Their love and support have been instrumental in my achievements. Special thanks also go to my extensive family, including my sisters, brothers and their beloved children. Throughout my PhD journey, these wonderful people provided me with the love, support, attention and prayers I needed to overcome the challenges I faced. My special thanks extend to my own small family, my husband and my son, Battal. I acknowledge that this work would not have been possible without their presence in my life.

On a practical note, I would also like to extend my thanks to all the members of the institutions that assisted and supported me in obtaining the data and information required to complete this project. These organizations include the Saudi Electricity Company, King Abdullah Petroleum Studies and Research Center (KAPSARC), Met Office National Meteorological Archive, Weathernews France, and Alfanar Construction. I am also deeply grateful to Lancaster University for providing me with an exceptional academic environment and abundant resources. Additionally, I appreciate Taif University for awarding me a scholarship to study in the United Kingdom in the field of networking and cybersecurity.

Contents

List of Figures

List of Tables

Nomenclature

ACF Autocorrelation function

AdaBoost Adaptive gradient boosting machines

AGNES Agglomerative nesting

AMI Advanced metering infrastructure

ANN Artificial neural network

AP Affinity propagation

AUC Area under the curve

BMS Building management system

CART Classification and regression tree

CatGBM Categorical gradient boosting machines

CFSFDP Clustering by fast search and finding density peaks

CNN Convolutional neural network

D-FFNN Dual deep feed forward neural network

DBSCAN Density-based spatial clustering of applications with noise

DC Direct current

DR Demand response

DRES Distributed renewable energy sources

DSO Distribution system operator

DT Decision tree

EMS Energy management system

FCM Fuzzy C-means clustering

FDI False data injection

FI Feature importance

FIT Feed in tariff

FPR False positive rate

GBM Gradient boosting machines

GPS Global positioning system

GRU Gated recurrent unit

HAN Home area network

HEMS Home energy management system

ICS Industrial control system

IEA International energy agency

IED Intelligent electronic device

K-NN K-nearest neighbours

LAN Local area network

LightGBM Light gradient boosting machines

LOF Local outlier factor

LSTM Long short term memory

MITM Man in the middle

MLPNN Multi-layer perceptron Neural Network

NAN Neighbourhood area network

OPF Optimum path forest

PACF Partial autocorrelation function

PCA Principal component analysis

PDC Phasor data concentrator

PMU Phasor measurement unit

POWER Predictions of worldwide energy resources

RBF Rradial basis function

RFE Recursive feature elimination

RFID Radio frequency identification

RNN Recurrent neural network

ROC Receiver operating characteristic

RTU Remote terminal unit

SCADA Supervisory control and data acquisition

SMOTE Synthetic minority over-sampling technique

SVM Support vector machine

T&D Transmission and distribution

TPR True positive rate

TSO Transmission system operator

UFS Univariate feature selection

WAMS Wide area measurement system

WAN Wide area network

XGBoost Extreme gradient boosting

Chapter 1

Introduction

Energy[1] is broadly regarded as a fundamental element contributing to both social well-being and sustainable development. Electrical energy grids have grown considerably with the shift from agrarian societies to industrial and information-based societies (Amin and Stringer, 2008). They have expanded in an unprecedented manner, reaching practically every residence, factory and institution in developed countries, while also rapidly expanding in the developing world.

Conventional electric grids, predating the smart grid paradigm, typically include a collection of independent and enormous current networks. They consist of thousands of central generation plants that produce electricity resources, such as uranium for nuclear power and coal as fossil fuel for traditional thermal power generation. Energy is transmitted from these central plants through high-voltage transmission networks to distributed load centres, and then delivered via low-voltage networks to energy consumers. This process is entirely centralised and managed by monopoly utility providers (Collier, 2017). Nevertheless, in order to contribute to the global net-zero initiative, the energy sector has sought to fully exploit a combination of distributed energy sources (DRES), nuclear generators and fossil-fuel facilities (Ekanayake et al., 2012). However, the management of this combination of energy resources is challenging in conventional grids. The monitoring and control requirements necessary for the cost-effective operation of such a system are unavailable in conventional grid underpinnings (Farhangi, 2009; Goudarzi et al., 2022).

This deficiency of conventional grids is being addressed by the development of next-generation energy grids, known as smart grids (Tuballa and Abundo, 2016). Smart grids provide an opportunity to leverage advanced information and communication technology and modernise electrical energy systems (Ekanayake et al., 2012). To put it simply, a smart grid is a cyber-physical system that comprises a cyber infrastructure (i.e. data-collection infrastructure) tightly integrated with a conventional energy

[1]This thesis focuses on energy delivered by electricity networks and not gas.

system to enable stakeholders to exchange both energy and information (Judge et al., 2022). By virtue of smart grids, data and measurements are continuously collected and processed at various levels of smart grid infrastructures; these are categorised as energy generation, transmission and distribution (T&D) and end-user. This data-collection infrastructure facilitates the monitoring, protection and control of energy systems effectively and reliably, and it offers significant opportunities for decarbonising the energy sector at a realistic cost (Ekanayake et al., 2012; Farhangi, 2009).

Nevertheless, the introduction of the smart grid paradigm has contributed to the expansion of security threats in energy systems. For instance, the infamous cyber-attack on Ukraine's energy sector resulted in power outages that affected around $225,000$ consumers for several hours (Lu et al., 2019). Another similar attack targeted the main electricity supplier in Johannesburg, South Africa; it caused major disruption of the electricity supply in some residential areas, leaving them without electricity (BBC News, 2019). In addition, there have been unconfirmed attempts at cyber-attacks on the national grid infrastructure of the US and the UK, wherein the potential hackers tried to break into the utilities' networks to disrupt their services (Sobczak, 2019; Pfeifer, Fildes, and Ram, 2018). Such a scenario could allow malicious entities to further orchestrate energy theft activities without being detected, causing major losses to the utility (McLaughlin, Holbert, et al., 2013; Mahmoud et al., 2020; Yan and H. Wen, 2021). In this regard, the diversity of hardware and software technologies employed within smart grids and the lack of holistic grid-specific security practices facilitate the development of new energy theft techniques (Jiang et al., 2014a; Yao et al., 2019; Qi et al., 2016).

Considering that the attack vectors underpinning energy theft span numerous smart-grid vulnerabilities, there is no single globally accepted definition of the energy theft threat. The intentional tapping of energy from distribution networks to physically steal energy is theft (Bihl and Hajjar, 2017; W. Han and Xiao, 2016; Weslowski, 1976). Tampering with metering systems to lower energy bills (e.g. (Z. Zheng et al., 2018; Jindal, Dua, et al., 2016; McLaughlin, Podkuiko, and McDaniel, 2009; Ngamchuen and Pirak, 2013; Yip, Wong, et al., 2017a; Sharma and Majumdar, 2020; Zanetti et al., 2017)) and the fraudulent maximisation of generated energy measurements (e.g. (Mahmoud et al., 2020; Shaaban et al., 2021; Yuan, M.-g. Shi, and Sun, 2015; Yuan, M. Shi, and Sun, 2015; Krishna, Gunter, and Sanders, 2018)) also constitute energy theft. Manipulating the measurements of the T&D system is also financially profitable misconduct in energy systems (Tajer, 2017; Xie, Mo, and Sinopoli, 2010; Goudarzi et al., 2022; G. Cheng et al., 2022). Therefore, we can define energy theft as the illegitimate exploitation of energy grid infrastructures, components, communication networks, applications and management systems for the purpose of manipulating the business model and gaining monetary profits.

Globally, energy theft has reportedly caused significant losses of electric energy for

utility providers. These non-technical losses amounted to £1.4 billion per annum in Brazil, and for a single energy provider in Canada, they resulted in an average annual loss of 850 GWh (£55 million of monetary loss) (Raggi et al., 2020). Each year in the UK alone, energy worth £400 million is stolen, leading to inflated customer bills (Yorukoglu et al., 2016). As reported in (Yao et al., 2019), energy theft causes utility companies worldwide to lose more than £19 billion annually. Regardless of whether such theft attacks are executed by a single consumer or on a large scale, the losses incurred by providers due to energy theft are undesirable and highly significant.

In addition to monetary and energy losses, the literature highlights a variety of other consequences of energy theft. For example, theft-related activities have a negative impact on the reliability of energy grids. Because of the energy lost in theft-related activities, the supply system may be overloaded by customer demand, causing its reliability to deteriorate (i.e. sometimes individuals experience power outages) (Rouzbahani, Karimipour, and Lei, 2020). Furthermore, when utility providers operate at a monetary loss, they must raise the entire billing amount, requiring legitimate users to pay more for their energy (Ahmed et al., 2022). Several studies conducted in 2019 indicated that almost 80% of $2,000$ UK residents were unaware that energy theft directly affected them (Safe, 2018). These studies also revealed that due to energy thefts, £20 per annum was added, on average, to household bills. Thus, millions of consumers pay for energy that they have not used and, most importantly, not stolen.

1.1 Problem statement

Energy providers around the world have reduced their monetary and energy losses by detecting energy theft. Conventional methods for detecting energy theft primarily rely on physical inspections of areas where energy thefts are anticipated to increase. However, these conventional methods are time-consuming, inaccurate, costly and labour-intensive, thereby decreasing the return on investment in anti-energy theft initiatives (Jindal, Dua, et al., 2016; Aldegheishem et al., 2021; Gao, Foggo, and Yu, 2019a). Advanced data-driven detection approaches fostered by machine learning techniques leverage the integrated cyber infrastructure of smart grids to enable these inspections to be conducted more effectively (Gao, Foggo, and Yu, 2019b; K. Zheng et al., 2018).

Nonetheless, the majority of these approaches have been limited in their detection scope due to their explicit focus on particular types of measurements or properties of the overall smart grid ecosystem. Hence, numerous schemes exist for detecting energy theft through consumption-related readings ((Messinis, Rigas, and Hatziargyriou, 2019a; Zanetti et al., 2017; Sharma and Majumdar, 2020; M. Wen et al., 2021;

Gunturi and Sarkar, 2021; Messinis, Rigas, and Hatziargyriou, 2019b; Yao et al., 2019; Z. Zheng et al., 2018; Gao, Foggo, and Yu, 2019b; Cody, Ford, and Siraj, 2015; Jindal, Dua, et al., 2016)), or T&D measurements (e.g., (Buzau et al., 2018; Ashrafuzzaman, Das, et al., 2020; Esmalifalak et al., 2017; Ying Zhang, J. Wang, and B. Chen, 2020; Mukherjee, Chakraborty, and Ghosh, 2022; Hegazy et al., 2022)). However, only a limited number of detection solutions focus on the detection of theft-related DRES generation measurements (Yuan, M. Shi, and Sun, 2015; Mahmoud et al., 2020; Shaaban et al., 2021; Krishna, Gunter, and Sanders, 2018).

Furthermore, the deployment and synchronisation of the aforementioned monolith in practice by providers at different levels of aggregation would be ineffective, as they pose highly demanding computational requirements. Moreover, the algorithmic properties involved in such approaches have been proven to be unable to sufficiently distinguish theft-related activities from anomalous events that could be caused by non-malicious intent (e.g. grid equipment misconfiguration) (Messinis and Hatziargyriou, 2018; Maamar and Benahmed, 2018). Finally, the vast majority of theft detection solutions fail to effectively adapt and re-optimise their detection thresholds; hence, consideration should be given to how these need to change in response to the addition of new types of grid components and the adoption of new technologies.

1.2 Motivation and objectives

Keeping these issues in mind, this thesis focuses on delivering a practical data-driven framework to address these challenges to a considerable degree, in order to help minimise the losses caused by energy theft activities. To achieve this ambitious aim, several algorithmic–technical objectives have been established, as listed below:

1. To investigate the impact of the introduction of the smart grid paradigm on the security of energy systems against energy theft attacks.

2. To identify and assess the weaknesses of current energy theft detection schemes.

3. To design and implement a theft detection framework utilising diverse sources of measurements and considering the highly distributed nature of DRES.

4. To design and implement a practical, holistic and adaptive system for energy theft detection in consumption and generation measurements.

To achieve the aims and objectives of this thesis, the following research questions need to be addressed:

1. How does the introduction of smart grids enable larger attack vectors that provide a basis for energy theft?

2. How do we design a data-driven framework for detecting energy theft?

The latter question can be divided into the following sub-questions:

a How can we leverage diverse sources of measurement to identify energy theft attacks in DRES?

b How do we devise a generic method to accurately detect energy theft in scalable smart grids?

1.3 Research contributions

To answer these questions, this thesis explicitly contributes to the wider research community by developing the following:

1. An investigation and survey of advances in energy theft from different perspectives of the smart grid ecosystem. It revolves around energy data manipulation, considering the three functions of demand, supply and generation. A variety of vulnerabilities enable adversaries to exploit grid infrastructure components, communication networks and managements systems, with the intention of achieving monetary benefit. This review provides an overview of the different types of energy theft attacks in smart grids. It reviews the latest research studies on attack strategies that enable energy theft, and it outlines their key findings. Moreover, it discusses existing energy theft detection schemes and summarises the outstanding challenges. This work serves as a first stop for both general audiences and domain specialists looking for information regarding energy theft in present-day smart grid systems and markets.

2. A generically applicable theoretical framework for a data and machine learning based energy theft detection process. In general, data-driven energy theft detection solutions in industrial applications rely on a variety of measurements collected by the data-collection infrastructure integrated into modern energy ecosystems. They leverage a wide range of techniques and algorithms from a broad spectrum of knowledge, machine learning being the most prevalent. Therefore, we intend to take a step beyond the current literature by developing a structured, generically applicable framework for energy theft detection in smart grids. The proposed theoretical framework reviews existing theories and serves as a road map for constructing a data-driven strategy for energy theft detection in different scenarios.

3. A generic SCADA-agnostic DRES power profiling scheme. In general, this DRES profiling system enables automated feature selection and the fine-tuning of machine-learning-based regression models, and it can also adapt to

diverse measurement feeds. Through a proof-of-concept study of wind turbine deployments, this work demonstrates that the proposed system can operate adequately by using freely available third-party weather measurements. Thus, it introduces a SCADA-agnostic approach that can sufficiently serve a range of envisaged smart grid applications, such as malicious actor detection.

4. A novel and low computational cost DRES-based energy theft detection approach. This approach builds upon the previous contribution to provide a predictive data-driven detection solution that considers weather dynamics. It goes a step beyond current solutions by removing any dependence on SCADA measurements and by largely focusing on third-party and freely available measurements. Thus, it intends to tailor theft detection accuracy based on the explicit properties of the generation of individual DRES deployments.

5. An adaptive energy theft detection system capable of distinguishing the properties of energy consumption and generation theft as opposed to possible misconfigurations caused by non-malicious intent. The proposed system self-adapts through a self-learning mode of operation that is continuously updated as new measurements become available. Thus, it promotes low computational costs and its architecture can be easily integrated with smart grid infrastructures to realize next-generation cross-batch energy theft detection schemes.

1.4 Thesis outline

Chapter 2 is a review of energy theft data-driven attack strategies and detection methods. By considering various operational and functional layers in modern smart grids, this chapter critically assesses how energy theft can be formulated. Moreover, this chapter provides an overview of the grid demand, supply and control chain, with a focus on energy theft and the associated security flaws that currently exist in the smart grid ecosystem. Different models for theft detection in smart grids are categorised. Finally, the chapter discusses various open issues in the scope of data-driven energy theft detection methods, and it suggests solutions in this field.
Chapter 3 proposes a generically applicable theoretical framework of data-driven energy theft detection methodologies in smart energy grids. Moreover, in order to present case studies of the application of the proposed theoretical model to generate data-driven theft detection methodologies, this chapter provides an overview of the data-driven detection approaches designed, implemented and evaluated in this thesis.
Chapter 4 describes a SCADA-agnostic approach that uses freely available weather measurements to explicitly profile and forecast power generation, as produced in real wind turbine deployments. For this purpose, the chapter leverages various machine

learning libraries to demonstrate the applicability of our system and further compares it with forecasting outputs obtained when using SCADA measurements.
Chapter 5 introduces a predictive data-driven, SCADA-agnostic energy theft detection approach explicit to DRES-based scenarios. This chapter comprehensive formalises a DRES-based theft attack model, and it further assesses the performance of our framework by using freely available third-party weather measurements and relating them to real solar and wind turbine deployments in Australia and France.
Chapter 6 proposes a self-learning system that can detect theft activity and misconfiguration events in consumption and generation measurements, while continuously learning from a stream of incoming measurements without human interference. In this chapter, a comprehensive formalisation of an adversary model that highlights the properties of smart meter misconfiguration and theft-related activities is provided. Furthermore, extensive data-driven experiments are conducted to demonstrate the proposed system's efficacy using real-world energy measurements.
Chapter 7 concludes the thesis by focusing on the contributions made to the proposed study domain. In addition, this chapter suggests future research directions to build on the work presented in this thesis.

Chapter 2

Energy Theft in Smart Grids: A Survey on Attack Strategies and Detection Methods

Cyber-physical attacks on power grids aiming explicitly at energy theft are the most prominent and they have been reported to cause significant financial as well as functional losses to energy utility companies at a global scale (Jindal, Dua, et al., 2016). Hence, energy theft attacks cause major concerns to both providers and consumers. In order to prevent such energy and revenue losses, utility companies typically conduct physical inspections in the locations where energy theft is due to intensify (Jindal, Dua, et al., 2016). Nonetheless, such conventional energy theft detection tracking is time-consuming, inaccurate, costly, and labour-intensive (Aldegheishem et al., 2021; Gao, Foggo, and Yu, 2019a). Therefore, to deploy more effective theft countermeasures, providers need to make use of the present electricity market driven by the need to collect and analyze data. The facilitation of data-driven operation drives utility providers to embed smart metering equipment in various levels of the electricity flow within smart grids (Angelos K Marnerides et al., 2014).

The entire life cycle of gathering energy data runs through smart grid infrastructures which are categorized into electricity generation, transmission and distribution (T&D), and end-user infrastructure. This data collection infrastructure leads to the emergence of an advanced line of detection method driven by measurement-based data providing opportunities to address energy theft. Data-driven detection is able to reduce the risk of lateral attacks leading to energy theft and recognize anomalous system behaviours arising from such events. Thus, reduce revenue losses for service utilities (Zanetti et al., 2017; Messinis, Rigas, and Hatziargyriou, 2019a).

Although, a variety of data-driven detection methods have been developed, malicious actors continue to discover innovative strategies in an attempt to perpetrate

energy thefts across smart-grid infrastructures (Aydin and Gungor, 2018). In this regard, the smart grid data measurements and monitoring infrastructure can pave the way for more approaches to fabricate next generation data-driven theft attacks, thus increasing relative energy and financial losses. Authors in (McLaughlin, Podkuiko, and McDaniel, 2009) and (Jiang et al., 2014b) review these data-driven theft attacks from the perspective of power-system communication-layer architectures, based on adversary strategies targeting the integrity of the power system by manipulating power demand data. Moreover, authors in (Messinis and Hatziargyriou, 2018; Saeed et al., 2020) and (Yan and H. Wen, 2021) provide an overview of energy theft detection, including features employed, methodology and procedures, evaluation metrics, and a comparison of the performance of each detection method. However, these surveys were not focused on energy theft and do not consider recent advances in modern smart grids, as the nature of vulnerabilities and threats related to energy theft are constantly changing due to the increasing intersection of power grids with Internet-enabled cyber-physical systems (Mahmoud et al., 2020).

Motivated by these observations, we investigate and survey the advances in energy theft from different perspectives within the smart grid ecosystem revolving around energy data manipulation from all the three functions of demand, supply, and generation. We explicitly contribute in the wider research community for modern energy grids by providing:

1. The first survey work covering the largest spectrum of attack strategies available in the literature used for carrying out energy theft in the modern electricity market.

2. Introduction of an energy theft categorisation model which provides a comprehensive perspective on defining energy theft from various smart-grid data flows.

3. A critical assessment of lessons learned from the application of various approaches presently used for detecting energy theft.

4. Recommendations for suggested solutions with respect to the design of energy theft detection schemes as tailored with an extensive analysis of open issues.

The remainder of this chapter is organized as follows: Section 2.1 focuses on the key infrastructures consisting the modern power grids such as to relate attack vectors associated with energy theft. Section 2.2 provides a comprehensive analysis on energy theft attacks. In Section 2.3, we categorize and discuss algorithms used in energy theft detection systems. Section 2.4 presents the existing gaps in research for energy theft detection and recommends suggested solutions. Finally, in Section 2.5, we conclude and summarise this chapter.

2.1 Smart grid components

Energy theft may span over multiple logical or physical entities and can be instrumented via numerous attack vectors affecting one or more of the systems consisting the modern smart grid. Within this work, the various properties of energy theft are discussed in terms of the intrinsic characteristics of each of these infrastructures. Therefore this section is dedicated at presenting an overview of the infrastructure of the smart grid with its core components.

One of the main goals within the modern smart grid is to ensure the optimal operation of the electricity supply chain. As shown in Fig. 2.1, the end-to-end energy supply chain is decomposed into three distinct phases; i) generation , ii) transmission and distribution (i.e. T&D) and, iii) end user consumption. All three phases are directly dependent on explicit technologies, administrative domains and networked power system infrastructures. Each of these entities pose unique vulnerabilities that can enable energy theft (Otuoze, Mustafa, and Larik, 2018; Jindal, Schaeffer-Filho, et al., 2020; Messinis, Rigas, and Hatziargyriou, 2019a; Shaaban et al., 2021).

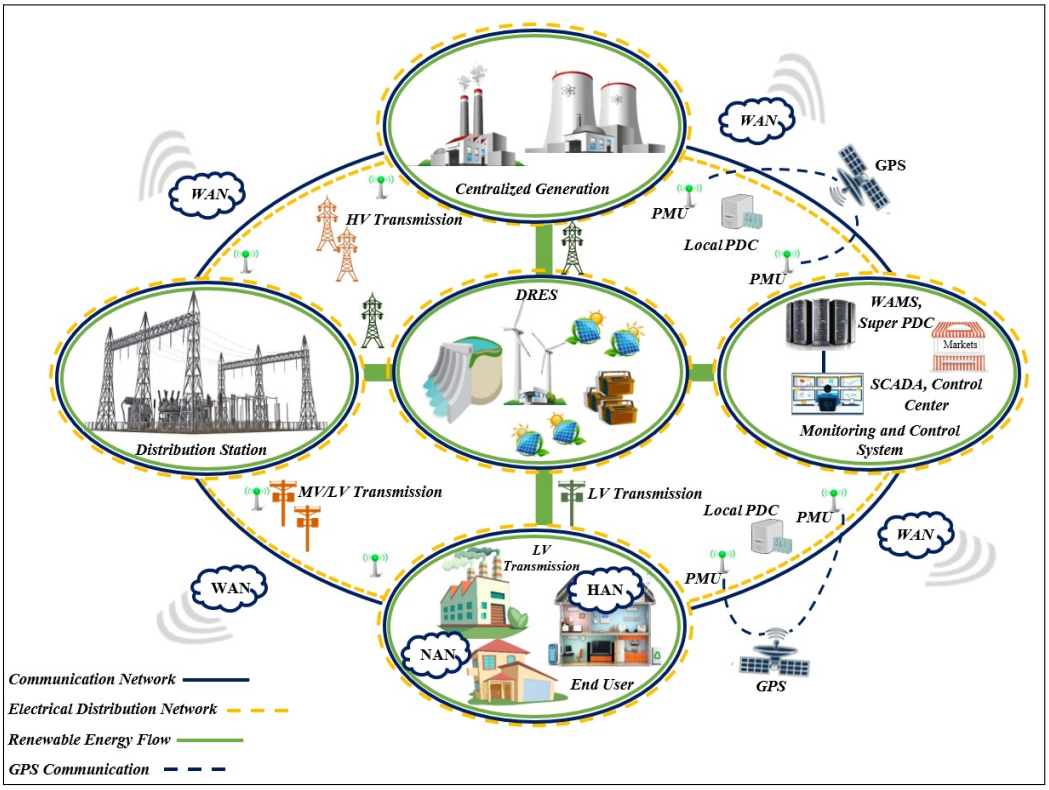

Figure 2.1: Phases and components of the energy supply chain in the smart grid.

The energy generation phase is achieved within large, centralised power stations that nowadays are interfaced with power generation DRES deployments and are commonly owned by the national transmission energy network controlled by one or a set of transmission system operators (TSOs). Each TSO is engaged through a competitive energy trading market scheme with a number of distribution system operators (DSOs) in order to supply them with electricity to be distributed to end-consumers [1]. DSOs may also have a direct interface and own DRES deployments or they frequently have an energy trading contract with end-consumers or third-party DRES owners that contribute directly in the energy generation phase.

In general, any control and management (sub)systems alongside the electro-mechanical set of power systems enabling data and energy flows spanning the energy supply chain are underpinned by diverse and ubiquitous data communication technologies. Fig. 2.1, indicatively illustrates a variety of potential networking technologies and deployment setups that could be employed in smart grids. Similarly with the energy trading market, the business model behind the ownership of these deployments depends on a number of aspects related to country-level legislation and policies (Burke and Stephens, 2018) and it is out of the interest within this thesis.

2.1.1 Energy generation

2.1.1.1 Centralised generation

Generation systems are categorised to operate either in a centralised or a decentralised fashion. Centralised generation produce large-scale electricity at power stations, utilising fossil fuels and nuclear plants or renewable resources such as hydroelectric power plants, wind and solar farms. These centralised systems are usually placed in remote areas that are distant from the end users. They are linked to distributed stations owned by a given DSO via a network of HV transmission lines operated by a TSO (EPA, 2018). The DSO stations are responsible for transmitting electricity through the medium and low-voltage grids to multiple end users (Yip, W.-N. Tan, et al., 2018).

2.1.1.2 Distributed Renewable Energy Sources (DRES)

DRES have evolved to act as an integral element of the electricity generation infrastructure aiding the needs of the backbone grid in terms of critical ancillary services (e.g., frequency regulation, reactive power) enabling grid stabilisation, diversifying energy trading and most importantly matching the peak during overloaded periods (C. Li and Shen, 2019; Shilay et al., 2017; Banshwar et al., 2017). Moreover,

[1] In the USA a TSO may be referred to as an independent system operator (ISO) and a DSO as a regional transmission operator (RTO).

DRES deployments are currently considered as the most suitable components for contributing towards the reduction of global carbon emissions (Shilay et al., 2017). According to the international energy agency (IEA), DRES deployments have contributed to 40% of the total primary energy supply globally in 2020 (IEA, 2020).

Energy generation billing and trading for DRES is currently achieved via two distinct systems; i) net metering and, ii) feed in tariffs (FIT). Net metering operates with a single meter and employs a model where prosumers use their own DRES-based generated power on-site and any surplus is considered as a future credit on their billing issued by their DSO. On the other hand, FIT operates based on two smart meters residing at the prosumer end dealing with the capturing of energy generation and consumption rates independently. By contrast to net metering, FIT decouples the monitoring process and facilitates a simpler data processing framework for energy trading as well as billing, thus it was extensively adopted in a number of developed countries such as the United Kingdom, Canada, Japan, China, and Australia (Mahmoud et al., 2020).

Despite of the various benefits offered by DRES deployments, their direct dependency on natural resources (e.g., wind, solar radiation) that are in some cases unpredictable to fully forecast may cause challenges and higher complexity within the overall grid optimisation process. Thus increasing risks related to aspects of management of congestion, regulation of voltage, and grid stability (B. Zhao et al., 2018; X. Han et al., 2018). In parallel, the integration of DRES involves diverse types of data communication and system-on-chip technologies that are commonly manufactured with minimal security (Bor et al., 2019; Jindal, Angelos K. Marnerides, et al., 2019). Hence, enlarging the spectrum of cyber attacks that could be initiated such as to support potential energy theft acts (Krishna, Gunter, and Sanders, 2018).

2.1.2 Energy Transmission & Distribution (T&D)

2.1.2.1 T&D energy flow

The T&D infrastructure is responsible for enabling the transmission of power and further distribution of electricity to the consumers. As depicted in Fig. 2.1, T&D infrastructures may be categorised into the low-voltage (LV), high voltage (HV) and medium voltage (MV) power networks. Throughout the years, the topology for these power networks has evolved from an ordinary radial structure to interconnected or consistent networks, which has guaranteed higher reliability, operational economy, and best equipment use. Primarily, the electricity produced by the centralized electricity generation systems is transported to different distribution stations over HV transmission lines, which is then supplied to the end users through the widespread transmission lines of MV and LV networks. In parallel, modern T&D

infrastructures also distribute energy generated at DRES deployments through MV-LV substations (C. Wang et al., 2017).

2.1.2.2 T&D data communication

The data communication network underpinning the operations of T&D infrastructures commonly consists of two types of networked deployments that interact with the end-consumer home area network (HAN). As demonstrated in Fig. 2.2, end-to-end data communication between the T&D infrastructure and a HAN is achieved via a wide area network (WAN) interacting with a set of neighbourhood area networks (NANs).

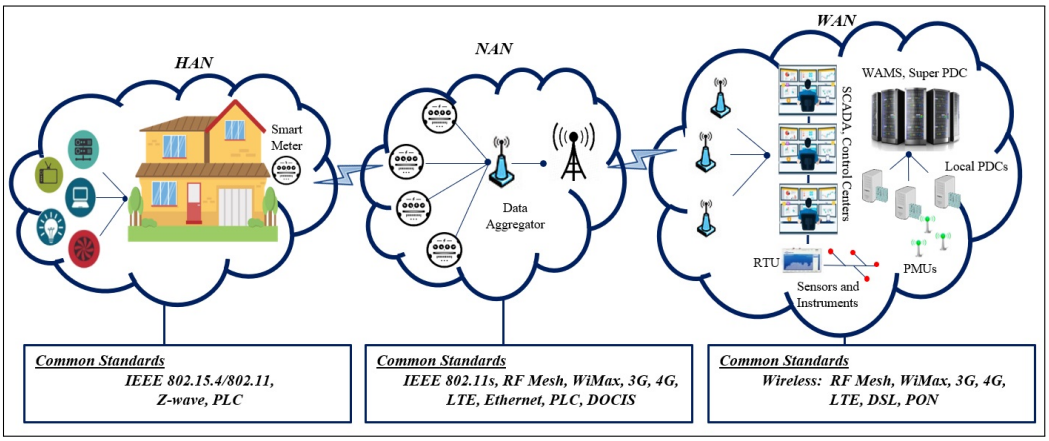

Figure 2.2: Exemplar smart grid network architecture highlighting some of the main data communication standards.

A WAN typically represents the aggregation of NANs and it is mapped at the scale of a city-wide network considering data flows related to energy distributed by multiple micro-grids where each micro-grid is linked with a particular NAN. In real deployments, the structure of a WAN is quite diverse since it may consist of multiple networking technologies with varying physical, logical and software components dealing with network control and management (Ogbodo, Dorrell, and Abu-Mahfouz, 2017; Yang Zhang, T. Huang, and Bompard, 2018). On the other hand, NANs can be considered as a subset of a WAN since they support smaller geographical regions and they act as proxies of a given WAN for functions related to connectivity and data aggregation of HANs with the main WAN. In general, a WAN or a set of WANs alongside related NANs and HANs are not necessarily always owned by corresponding TSOs or DSOs as they could be managed and maintained by third-party network providers (e.g., Internet Service Providers) or community entities (e.g., municipality).

2.1.2.3 Data acquisition & management

The actual interface of data communication with data-driven control and management of the processes explicit to reliable and resilience distribution of energy is achieved via network-enabled cyber-physical systems such as supervisory control and data acquisition (SCADA) systems. These systems are nowadays the most frequently used systems within modern T&D infrastructures. SCADA systems provide native integration of data communication technologies and system components such as remote terminal units (RTUs) and intelligent electronic devices (IEDs) (T. Liu et al., 2017; S. Tan et al., 2017). The data communication reliability offered by SCADA systems enables TSO/DSO control centres to develop close to real-time state estimation algorithms in order to optimise the grid's performance and increase situation-awareness (Sundararajan et al., 2019; Prado et al., 2019).

A relatively recent alternative approach to SCADA are wide area measurement systems (WAMS) (RB and GM, 2015). WAMS are embedded with new data acquisition technologies facilitating synchronised measurements between remote T&D deployments (e.g, micro-grids, substations) and facilitate the basis for monitoring, operation and control (Rezaee and Moghaddam, 2019). In practise, WAMS may be decomposed by a set of distributed Phasor measurement units (PMUs) and phasor data concentrators (PDCs) that sample data related to the waveform and the analog voltage of remote sites through a global positioning system (GPS) clock (RB and GM, 2015; Tian and Sansavini, 2016).

2.1.3 End user infrastructure

2.1.3.1 Advanced Metering Infrastructure (AMI)

AMIs are considered one of the fundamental components within the smart functionalities of the smart grid. The operation of such infrastructures achieves end-to-end metering in order to support the billing and trading processes between an end-consumer or prosumer and a DSO/TSO. A core innovation behind AMIs lies with the integration of smart meters within residential households or business buildings. In most developed and many of the developing countries, smart meters have replaced the traditional mechanical and analogue meters and they enable various services. Apart from the real-time logging of measurements related to end user energy consumption (i.e. demand data), smart meters also assess other features such as voltage levels as well as real-time monitoring (Lighari, Jensen, Shaikh, et al., 2014).

As already mentioned, data captured by smart meters contribute to the overall demand response (DR) model and they are transmitted through low-powered communication and automation protocols (e.g., ZigBee, Z-Wave) in synergy with upper layer application protocols (e.g., HTTP/HTTPS) supported by their corresponding

HAN. Fig. 2.2 provides an exemplar illustration in which smart meter measurements are locally aggregated within a HAN and are further distributed to the corresponding T&D infrastructure through an adjacent NAN interacting with a WAN. The sampling rate for measurements gathered by individual smart meters falls with a pre-defined schedule agreed between the end-consumer or prosumer with its corresponding DSO. Normally, measurements are agreed to be sent in $5, 15, 30$, or 60 minute intervals (Mohammad, 2018; Lighari, Jensen, Shaikh, et al., 2014; Ghosal and Conti, 2019; Ikpehai, Adebisi, and Rabie, 2016).

2.1.3.2 Energy Management System (EMS)

The adequate management and reactive control of energy usage and production in end user deployments is achieved through the installation of EMS instances. Such instances may be directly interfacing with a given DSO or through proxy third-party stakeholders maintaining and supporting large-scale EMS deployments. From the end user perspective, there is a variety of EMS types coming with specific functionalities such as home energy management systems (HEMSs) and building management systems (BMSs)[2]. In parallel, EMS can also be present at a larger scale deployed either at a centralised or a distributed topology aggregating measurements for the T&D insfrastructure (Arcos-Aviles et al., 2016; Solanki, Bhattacharya, and Canizares, 2017; Venayagamoorthy et al., 2016). Nonetheless, the main role of an EMS instance at the end user infrastructure is to optimise energy consumption for an individual or a set of individuals through controlling the various appliances residing within a given building or household (Jindal, Bhambu, et al., 2020). Hence, EMS software instances are usually composed of a controller instructed by advanced energy optimisation algorithmic components coupled with rule-based control functions orchestrating the operations of appliances (Angelos K Marnerides et al., 2014; Din, Mauthe, and Angelos K Marnerides, 2018).

2.1.4 Grid effectiveness pillars

The effectiveness of the grid in all levels depends on the performance of both quantitative as well as qualitative indicators. For instance, the reliable operation of the energy grid directly affects the well-being and safety of consumers whereas well-being is not a fully quantifiable parameter and, in parallel, grid reliability depends on quantifiable performance metrics (e.g., demand-supply rate) (Shokoya and Raji, 2019). Moreover, cyber-physical challenges, such as attacks enabling energy theft may affect directly grid optimisation processes, thus impacting grid reliability with a cascading impact over user safety since some power system machinery could be

[2]Discussion of EMS variations is out of scope for this chapter.

affected and malfunctioning (S. A. Salinas and P. Li, 2016; Czechowski and Kosek, 2016). The latter example has a number of parameters that are not necessarily quantifiable (e.g., grid security level, safety impact on consumers/prosumers), hence a holistic correlation scheme between the aforementioned pillars is an extremely challenging task.

As evidenced in Fig. 2.3 this work relates grid effectiveness with the three broad domains of reliability, resilience and safety that we refer to as pillars. We exploit definitions developed throughout the years and summarise the definitions of the three inter-related pillars in order to structurally assess the energy theft impact in the overall grid effectiveness (Kawann, 2002; Jufri, Widiputra, and Jung, 2019; Czechowski and Kosek, 2016):

1. Grid reliability: preservation of continuous energy supply to end consumers.
2. Grid resilience: preservation of continuous energy supply to end users with an acceptable level of energy quality while under stress or faults.
3. User safety: ensure that an individual or a group of individuals utilising or maintaining the grid and its services are not physically affected.

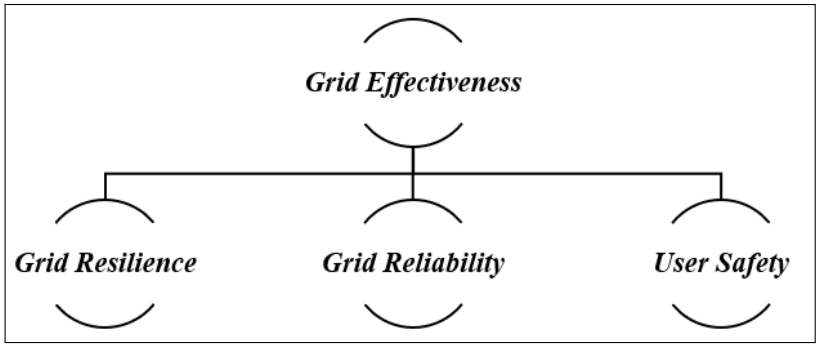

Figure 2.3: Grid effectiveness pillars.

This survey acknowledges that the highlighted pillars are considered widely as independent research domains themselves. Hence, deeper investigation on the structure and properties of these pillars is out of the context within this chapter.

2.2 Energy theft

Energy theft can be broadly defined as a case where individuals manipulate their consumption measurements in various ways, such as physically tapping (hooking) into

distribution lines or paying less than they should due to tampering with or bypassing their meters (Czechowski and Kosek, 2016). However, attack vectors underpinning energy theft span numerous vulnerability domains due to the emergence of a plethora of smart grid applications (e.g., energy trading) that rely on inherently vulnerable networked environments as a result of the convergence of diverse legacy power systems with Internet technologies (e.g., ICS deployments, metering). In general, energy theft can be instrumented through a variety of techniques exploiting both physical as well as data or communication-oriented properties of the current grid (Bihl and Hajjar, 2017; Smith, 2004; Jindal, Dua, et al., 2016; Mahmoud et al., 2020). Hence, the adequate categorisation of energy theft types is a highly challenging task.

In order to address the aforementioned challenge and appropriately structure the focus within this work, we identify two distinct classes of energy theft:

1. *Data-agnostic energy theft*: the act of physical tampering of power components through techniques such as obstruction of electro-mechanical meters, cable tapping, bypassing meters as well as energy harvesting.

2. *Data-driven energy theft*: the act of manipulating and altering communication and/or energy measurement-related data generated and/or logged at any networked metering (e.g., smart-meter), management (e.g., SCADA system) and control device (e.g., PLC) as well as billing software (e.g., utility mobile apps) aiming at reporting false energy information to the power distribution authority (e.g., a DSO).

Both classes target either of the bidirectional energy or data flow between different grid aggregation points (e.g., T&D, end user, generation) and they have seen a considerable level of attention from the research community as well as the society in general (Mahmoud et al., 2020; Jindal, Schaeffer-Filho, et al., 2020). Moreover, both types have shown to be applicable in all three levels of aggregation within a smart grid. Hence, energy theft can be deployed in the power generation infrastructure, the T&D network as well as the end-consumer level.

This work argues that the main concept of a given theft attack can be abstracted by a discrete function in which inter-dependent variables are tailored based on the targeted infrastructures composing a complete smart grid deployment. Hence, the function may vary depending on the variable-specific adjustments conducted by a malicious actor based on the intrinsic properties of a given smart grid (sub)infrastructure (e.g., communication, power). Commonly, malicious actors attempt to target a set of diverse vulnerabilities of both system and network components from all three infrastructures described herein. Evidently, energy theft in all three infrastructures has considerably increased due to the data-oriented functioning of the business layer as envisaged in the current smart grid reference architectures (e.g., SGAM (CEN-CENELEC-ETSI, 2012)).

2.2.1 Energy theft model

Energy theft in the context of the smart grid can be abstracted using various generalised approaches such as (Mahmoud et al., 2020; Esmalifalak et al., 2017; Punmiya and Choe, 2019). We indicate ways in which energy theft can be modeled from the perspective of manipulating generation, supply and demand data respectively. The proposed approaches rely on the notation denoted in Table 2.1.

Table 2.1: Energy theft model notation.

Ec	Demand node energy consumption
Er	Prosumer node energy generation
Es	Energy supply by T&D control nodes
NTL	Cumulative non-technical energy loss
TL	Technical energy loss
G	T&D grid
S	Number of grid supply nodes
M	Number of energy distribution buses
N	Number of total nodes
P	Number of prosumer nodes
Q	Number of consumer nodes
α	Theft coefficient on generation data
β	Theft coefficient on supply data
γ	Theft coefficient on demand data

As depicted in Fig. 2.4, we consider a grid G in a NAN to be defined by a set of N connected nodes and M connecting energy buses. A node is indicated as a prosumer node if it has a local DRES; otherwise, the node is indicated as a demand node.

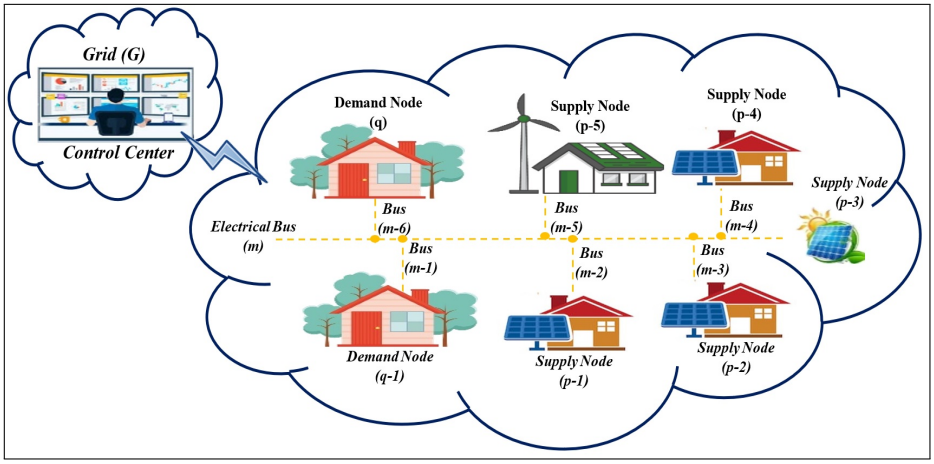

Figure 2.4: Energy grid model consisting of supply and demand nodes.

Let $TL_j(t)$ denote technical energy losses caused by wires and equipment resistance under the normal, theft-free condition in the j^{th} bus, where $j \in M$. We also consider the cumulative non-technical energy loss, NTL in G expressed as:

$$NTL(t) = \sum_{i=1}^{S} Es_i(t) + \sum_{k=1}^{P} Er_k(t) - \sum_{h=1}^{Q} Ec_h(t) + \sum_{j=1}^{M} TL_j(t) \tag{2.1}$$

Here, we characterise NTL in energy system G as it is essential to understanding the overall energy dynamics within the system. It can be broken down into four primary elements: the first two terms in Equation 2.1 represent the total energy supplied to the system. The summation $\sum_{i=1}^{S} Es_i(t)$ represents the energy supplied via T&D supply nodes, where $i \in S$. These nodes are responsible for delivering electrical energy from various sources owned by providers. The following term, $\sum_{k=1}^{P} Er_k(t)$, represents the total energy generated within G using DRES owned by presumers, where $k \in P$. The third term in Equation 2.1, $-\sum_{h=1}^{Q} Ec_h(t)$, is the total energy consumed by consumer nodes, where $h \in Q$. Consumer nodes represent residential, commercial and industrial consumers of energy within G. The final term, $\sum_{j=1}^{M} TL_j(t)$, aggregates the technical losses incurred during energy transmission over buses $j \in M$. These losses occur due to inherent resistance in wires and equipment and are an intrinsic aspect of energy transmission. The range of values for these technical losses is well-documented in the literature and is typically between 6% and 8% of the energy transmitted via the T&D infrastructure (EIA, 2016).

2.2.1.1 Generation data-oriented theft

We consider that various data manipulation attacks may be conducted on DRES generation data (Yuan, M. Shi, and Sun, 2015) and two-metering end-user deployments (Mahmoud et al., 2020) on the prosumer site. Alongside the inability to accurately predict weather fluctuations affecting energy generation, we abstract the total electrical energy injected to the power grid by the $k \in P$ supply nodes during an energy theft attempt to be:

$$\sum_{k=1}^{P} Er_k(t) = \sum_{k=1}^{P} \alpha_k Er_k(t) \tag{2.2}$$

where $\alpha_k(t) \in \mathbb{R}^+$ is the theft coefficient for each supply node and two outcomes for this coefficient are possible being:

$$\begin{cases} 1 < \alpha_k(t) < \infty, & \textit{malicious prosumers} \\ \alpha_k(t) = 1, & \textit{honest prosumers} \end{cases}$$

Each supply node $k \in P$ has a theft coefficient α at time t. In the legitimate case where no attack is present, the theft coefficient $\alpha_k(t)$ equals 1; meaning that

there are no discrepancies in the DRES generation measurement at node k, since $Er_k(t) = \alpha_k Er_k(t)$. However, in the generation data-oriented theft scenarios, the DRES generation measurements entailed within $Er_k(t)$ are scaled by an attacker based on an arbitrarily selected percentage, represented by $\alpha_k(t)$. For instance, the attacker in such a scenario may report 200% of the actual measurements when $\alpha_k(t) = 2$. Hence, we abstract malicious prosumers that report falsified metering for their DRES generation process. Consequently, the non-technical energy loss, NTL, will be greater or equal than that for the normal case (i.e. equation 2.1); since $\sum_{k=1}^{P} \alpha_k Er_k(t) \geq \sum_{k=1}^{P} Er_k(t)$.

2.2.1.2 Supply data-oriented theft

Let assume the generalised direct current (DC) model described in (Yao Liu, Ning, and Reiter, 2011; Esmalifalak et al., 2017) such as the energy supply in our grid G by S supply nodes to be defined as:

$$\sum_{i=1}^{S} Es_i(t) = \mathbf{J}\left(\sum_{j=1}^{M} \theta_j(t)\right) + \sum_{i=1}^{S} e_i(t) \tag{2.3}$$

where $\mathbf{J}\left(\sum_{j=1}^{M} \theta_j(t)\right)$ are the state variables composed of the voltages phase angles within a Jacobian matrix $\mathbf{J}$ and $\sum_{i=1}^{S} e_i(t)$ is the measurement error from supply nodes assumed to adhere to Gaussian noise e.

In energy theft, malicious actors normally manipulate a subset of measurement data to alter metering. Hence, the aggregation of energy supply Es from all supply nodes can be defined as:

$$\sum_{i=1}^{S} Es_i(t) = \mathbf{J}\left(\sum_{j=1}^{M} \theta_j(t)\right) + \sum_{i=1}^{S} e_i(t) + \beta_i(t) \tag{2.4}$$

where $\beta_i(t)$ is a vector representing maliciously injected data within the legitimate measurements captured by a given T&D control center. Essentially, $\beta_i(t)$ can be mapped as a False Data Injection (FDI) attack instrumented at various levels (e.g., communication protocol, metering protocol etc.).

2.2.1.3 Demand data-oriented theft

Consumers and/or prosumers are also capable to lie on their demand data by utilising FDI techniques to cause under-reporting of energy consumption (Kim et al., 2019; S. K. Singh, Bose, and Joshi, 2019; Gao, Foggo, and Yu, 2019b; Sharma and Majumdar, 2020; Bor et al., 2019). We denote as $\gamma_i(t)$ to be the theft coefficient

of node i at time t. Considering a demand data-oriented theft the non-technical loss NTL can be represented as (Punmiya and Choe, 2019):

$$NTL = \sum_{i=1}^{Q} \gamma_i Ec_i(t) \tag{2.5}$$

In this case, the NTL should be greater than that for the normal case; since $\sum_{i=1}^{Q} \gamma_i Ec_i(t) < \sum_{i=1}^{Q} Ec_i(t)$. Hence, the two possibilities for $\gamma_i(t)$ would be:

$$\begin{cases} 0 \leq \gamma_i(t) < 1, & malicious\, consumer/prosumer \\ \gamma_i(t) = 1, & honest\, consumer/prosumer \end{cases}$$

In more detail, each consumer and/or prosumer $i \in Q$ has a theft coefficient γ at time t. In the legitimate case assuming no attack enabling energy theft, there are no discrepancies in the demand measurements denoted by $Ec_i(t)$, since the relative theft coefficients $\gamma_i(t) = 1$ and $Ec_i(t) = \gamma_i(t)Ec_i(t)$. However, in the demand data-oriented theft, the attacker manipulates the demand measurement signal Ec_i at time t by enforcing an arbitrarily selected percentage entailed within $\gamma_i(t)$. Therefore, the attacker under reports demand measurements and just reports a small portion of measurements on a regular basis. For instance, an attacker could potentially report 50% of the actual demand data, when $\gamma = 0.5$.

2.2.2 Data-agnostic energy theft strategies

The most prevalent approach to data-agnostic theft is the direct tapping of LV infrastructures to consume free energy. In order to steal electricity, these attacks entail constructing an unauthorised overhead or buried line connection to the distribution transformer's line side. With this approach, a property or appliance that was not previously connected to the power grid can be connected. In the US, tapping has been used to steal energy since at least the 1890s (Bihl and Hajjar, 2017). Another data-agnostic energy theft strategy to reduce demand measurements is bypassing metering systems. It is an extension of the notion of tapping; however, in this case, prior electric energy is obtained by directly connecting the property wiring to the wires entering the meter wiring (Bihl and Hajjar, 2017). This theft strategy can either completely disconnect the metre or leave it connected with a bypass so that the metre continues to record some consumption (W. Han and Xiao, 2016).

For non-smart metres, a data-agnostic technique for energy theft involves attempting to interrupt meter measurements. This practice prevents metres from properly measuring consumption readings. Various mechanisms can be used here, such as reducing the counting wheel's speed by influencing a metre with an effective

electromagnetic field and preventing the counting wheel from moving by inserting photographic film between a metre's back housing and its glass front cover or by drilling a discrete hole in the metre's housing (Czechowski and Kosek, 2016). In solar energy generation, the use of a solar array simulator is a well-defined data-agnostic strategy for energy theft. It is capable of mimicking the output characteristics of the vast majority of photovoltaic cells (Yuan, M.-g. Shi, and Sun, 2015). A solar array simulator and a real solar energy generator are connected in parallel, with their respective outputs feeding the generation metre. Hence, the value of the generation metre exceeds the value of the generation from the actual generator, and in the worst-case scenario, fraudulent prosumers can obtain monetary gains without even installing a solar energy generator (Yuan, M. Shi, and Sun, 2015).

With energy harvesting methods, all voltage levels including high, medium and low networks, are vulnerable to data-agnostic energy theft attacks. There are legitimate instances of this activity, such as legitimately powering smart-grid sensors (Chang et al., 2012; Cetinkaya and Akan, 2017b; Cetinkaya and Akan, 2017a; Ozger, Cetinkaya, and Akan, 2018), and perhaps both legal and illegal instances of powering personal equipment without utility approval (Bihl and Hajjar, 2017). Siegel has just devised a technique for illegally harvesting energy from T&D infrastructure. Siegel has received numerous accolades for this method, including one from the Bremen University of the Arts (Bihl and Hajjar, 2017). These devices are revealed as free electricity sources and are widely regarded as a technologically advanced innovation; it is also anticipated that they will be widely utilised (Siegel, 2012; dansie, 2013; Moghe et al., 2009). However, this method of energy consumption may be viewed as an unauthorised method and an attack related to energy theft. It can strain the T&D system of any given utility. Although it is doubtful that a large amount of energy would be taken by a single energy harvesting device designed by Siegel, the total amount of energy stolen could be substantial in the future if these devices become more widespread (Bihl and Hajjar, 2017).

In view of the discussed data-agnostic attack vectors that may be employed to steal energy, this research categorises the different attack approaches in Table 2.2. As shown, these attacks may be conducted at any level of aggregation and can use a diverse range of resources (e.g., LV networks, and PV panels). Intriguingly, the stated pillars of grid efficiency indicated in Section 2.1 are vulnerable to a variety of data-agnostic theft attacks.

2.2.3 Data-driven energy theft strategies

Data-driven energy theft is orchestrated either through targeted or random methods (Jinping Hao et al., 2015; Shilay et al., 2017). Targeted theft refers to instances in which a malicious actor has full awareness of the vulnerability spectrum for a given

Table 2.2: Overview of the data-agnostic energy theft attacks.

Ref.	Strategies	Infrastructure	Resource	Attack Effect	Remarks
(Bihl and Hajjar, 2017)	Consumption meter manipulation, Direct stealing from T&D systems	Consumption T&D	Electro-mechanical meter, T&D networks	User safety and grid reliability	Explains thefts in relation to sophisticated energy applications, such as energy harvesting and the smart grid.
(Yuan, M.-g. Shi, and Sun, 2015; Yuan, M. Shi, and Sun, 2015)	Generation meter manipulation	Generation	PVs	Grid reliability	Introduces physical attack applied to inject energy into PV power systems which make the generation meter reading larger than normal.
(Czechowski and Kosek, 2016)	Consumption meter interruption, Direct stealing from T&D systems	Consumption T&D	Electro-mechanical meters	User safety	Assumes that the majority of electric energy theft is not perpetrated by wealthy and educated individuals, but by those with low to moderate wealth.
(Moghe et al., 2009; Chang et al., 2012; Cetinkaya and Akan, 2017b; Cetinkaya and Akan, 2017a; Ozger, Cetinkaya, and Akan, 2018)	Direct stealing from T&D systems	T&D	T&D networks	Grid reliability	Investigates the present technology for energy harvesting.

system consisting of a node set (e.g., DRES deployment), and purposely injects data such as to compromise its operation. Random methods usually refer to scenarios where a malicious actor disturbs the operation of individual nodes (e.g., a single DRES) by randomly flooding the application protocol dealing with metering data or by injecting corrupted measurement values while a node communicates with a centralised monitoring component (e.g., a SCADA system). Generally, theft triggered by random methods is detected with higher precision (Lore, Shila, and L. Ren, 2018).

Both targeted or random methods for energy theft may be triggered by a number of cyber-physical attack techniques. The most common technique employed in the context of energy theft is the combination of man in the middle (MITM) with false data injection (FDI) (El Mrabet et al., 2018). These attempts refer to cases where an individual with malicious intent intercepts and redirects communication traffic between a smart meter and an energy monitoring entity (e.g., SCADA instance in

a NAN) to its own hardware. Traffic is redirected to the malicious actor such as to modify legitimate measurements and further inject falsifying metering information and re-transmit it to the monitoring component in order to affect the energy billing process. Regardless of the attack scenario underpinning energy theft, there are always some necessary steps to be undertaken by a malicious actor. Fig. 2.5 briefly provides some core steps that are frequently practised.

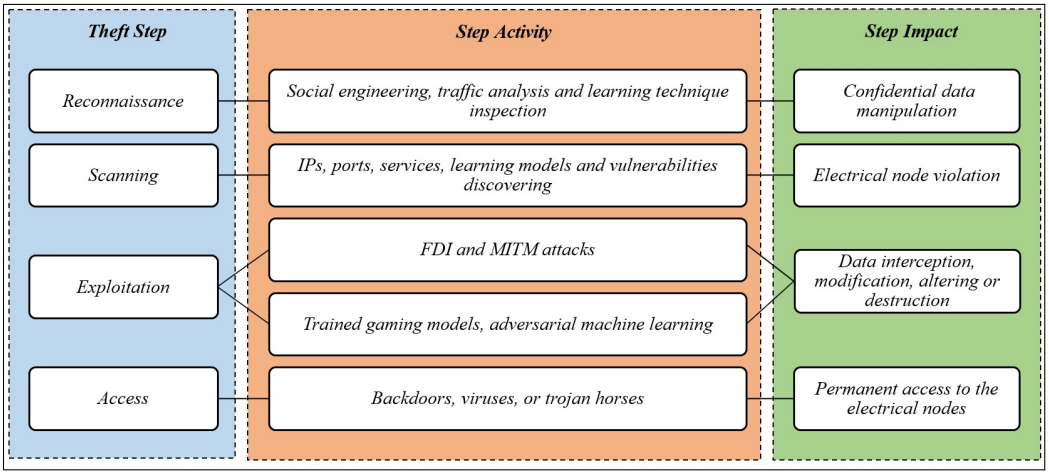

Figure 2.5: Steps and associated activities in cyber-physical attacks enabling energy theft.

We highlight four steps that in many cases are used concurrently in a given attack; i) reconnaissance, ii) scanning, iii) exploitation, and iv) access. Hence, there exists a number of variations of how the aforementioned synergistic use of MITM and FDI can be instrumented (C. Peng et al., 2019; Engebretson, 2013; El Mrabet et al., 2018). For instance, malicious actors could intercept general traffic at specific data recording entities (e.g., microgrid backend server) that they were aware of due to either scanning or reconnaissance such as to jeopardise the final data writing process with crafted, falsified measurements (McLaughlin, Podkuiko, and McDaniel, 2009).

Other examples, include a combination of physical tampering of meters at various power grid levels (e.g., T&D, end user smart-meters) where an attacker could identify through simple social engineering and bypassing of authentication protocols through ANSI optical ports with software such as Terminator that enables access (Mahmoud et al., 2020). In parallel, sophisticated MITM and FDI techniques may also consider the overall topology of a given grid deployment (Xie, Mo, and Sinopoli, 2010) in order to bypass any detection mechanisms whereas other utilise adversarial machine learning in order to game optimisation, scheduling and control processes within the EMS (Shilay et al., 2017; Bor et al., 2019). The aforementioned technique is relatively new and exploits the deficiencies of automated management functions by manipulating

and crafting falsified training data to machine learning-based processes that profile several measurements (e.g., ramp rate, power factor, reactive power) (Bor et al., 2019).

Given the diversity of the cyber-physical attack vectors enabling energy theft (Jindal, Angelos K. Marnerides, et al., 2019; Jindal, Schaeffer-Filho, et al., 2020), this work organises the various attack strategies based on their instrumentation and further impact in Table 2.3 and Table 2.4. As depicted, there has been a large volume in literature identifying, studying and further demonstrating that such attacks can be initiated at various aggregation levels by utilising different types of resources (e.g., SCADA, PV panels). Interestingly different types of attacks affect explicit grid efficiency pillars that we introduced in Section 2.1.

Table 2.3: Overview of the data-driven energy theft attacks.

Ref.	Strategies	Infrastructure	Resource	Attack Effect	Remarks
(Mahmoud et al., 2020)	Generation meter manipulation	Generation	PVs	Grid reliability	Introduces attack functions applied to manipulate the reported energy generation profile of PV power systems.
(Shaaban et al., 2021)	Generation meter manipulation	Generation	PVs	Grid reliability	Synthesizes a new presumer behaviour in an attack function (adding a fixed attack coefficient to the actual generation).
(Krishna, Gunter, and Sanders, 2018)	Generation meter manipulation	Generation	PVs and Wind turbines	Grid reliability	Assumes attacks manipulating the average of net generation while the detection mechanism is perceptible.
(Sundararajan et al., 2019)	Monitoring and control systems manipulation	T&D	PMU	Grid resilience	Summarizes different methods applied to commit data-driven theft against the grid measurements through WAMS manipulation.
(Pal, Sikdar, and Chow, 2017)	Monitoring and control systems manipulation	T&D	PMU and PDC	Grid resilience	Assumes attackers compromise one or more of the PMUs, PDCs, communication links or/and routers.
(Ashok, Govindarasu, and Ajjarapu, 2018)	Monitoring and control systems manipulation	T&D	SCADA	Grid resilience	Makes various assumptions about the attacks in the context of the current security mechanisms in SCADA networks.
(Punmiya and Choe, 2019)	Consumption meter manipulation	End user	Smart meter	User safety	Generates and labels real-time attack patterns for use with supervised detection algorithms.

Table 2.4: Overview of the data-driven energy theft attacks (Con.).

Ref.	Strategies	Infrastructure	Resource	Attack Effect	Remarks
(Basumallik et al., 2017)	Monitoring and control systems manipulation	T&D	PMU	Grid resilience	Assumes the attacker has access to only the PMU measurements at buses where the PMU has been compromised.
(S. K. Singh, Khanna, et al., 2017)	Monitoring and control systems manipulation	T&D	SCADA and PMU	Grid resilience	Assumes the attacker only compromises a single state variable. The attacker alters all the measurements to project the desired changed state variable.
(Xie, Mo, and Sinopoli, 2010)	Monitoring and control systems manipulation	T&D	SCADA	Grid resilience	Assumes the attacker can access several SCADA's sensors to compromise several measurements.
(Tajer, 2017)	Monitoring and control systems manipulation	T&D	SCADA	Grid resilience	Introduces a more realistic attack where the attackers have only inaccurate and incomplete information because of their restricted access to the grid.
(K. Zheng et al., 2018)	Consumption meter manipulation	End user	Smart meter	User safety	Introduces data-driven attacks enabling time-variant modifications on load profiles of the end users.
(Kim et al., 2019)	Consumption meter manipulation	End user	Smart meter	User safety	Models the energy loss resulting from meter manipulating, meter malfunctioning, and illegal bypassing.
(S. K. Singh, Bose, and Joshi, 2019)	Consumption data manipulation	End user	Smart meter	User safety	Introduces theft attacks based on the manipulation of appliance load profiles and the exclusion of heavy appliances from the actual measurements.
(Gao, Foggo, and Yu, 2019b)	Consumption meter manipulation	End user	Smart meter	User safety	Presents theft attack assuming the customer has DRES installation.
(Sharma and Majumdar, 2020)	Consumption meter manipulation	End user	Smart meter	User safety	Introduces a theft attack designed by a fraudulent employee, as such an employee fabricates the energy consumption readings based on past measurements, rather than reading the current actual measurements from the smart meter.

2.3 Energy theft detection methods

As briefly discussed in Section 2.2, energy theft can be *data-agnostic* and resulted purely from physical tampering of various grid components, or *data-driven* via manipulating, destroying or corrupting software processes with the goal to modify any data related to energy demand, generation or consumption. Throughout the years, both the industry and the research community have developed and employed techniques in aiming to detect any energy theft-related activities. In general, energy theft detection methods are structured under two main categories; i) hardware-based detection and ii) data-driven detection.

2.3.1 Hardware-based detection methods

Hardware-based solutions emphasise the installation of specialised metering equipment in order to identify and/or prevent electricity theft. For example, in (Khoo and Y. Cheng, 2011), the authors propose using radio-frequency identification (RFID) technology to assist energy providers in detecting energy theft. A case study demonstrates that a Chinese supplier who installed an RFID system obtained a positive return on their investment: more than $14,000$ dollars in savings when the RFID technology was implemented. Using a single-chip solution, the proposed approach of (Ngamchuen and Pirak, 2013) applies a smart anti-tampering algorithm for a single-phase smart meter. The findings demonstrate that the proposed approach provides a wide dynamic range of theft event detection, low false detection probability and rapid tampering detection. Moreover, the proposed algorithm still yields accurate tampering detection capabilities in the case of small energy measurements. Moreover, a study by (Dineshkumar, Ramanathan, and Ramasamy, 2015) presents an automatic meter-reading processor-based detecting system. When such a meter identifies a theft attempt, it sends a signal to the supplier via a global system for mobile networks. The results indicate that the proposed system is more efficient than 90%.

In general, hardware-based energy theft detection methods have numerous drawbacks, including the expense of deploying equipment, their vulnerability to failure and the complexity of maintaining devices. These constraints limit the development of these detection methods and have led to the emergence of data-driven approaches (Z. Zheng et al., 2018).

2.3.2 Data-driven detection methods

Generally, the data-driven energy theft detection is achieved through the algorithmic solution composition focusing on deviations of data related to aspects such as metering and billing. Hence, such detection schemes place a strong emphasis on analysing

data patterns through a variety of statistical tools and the majority utilises machine learning techniques. This chapter stratifies and discusses data-driven energy theft detection with respect to three main categories; i) classification-based, ii) regression-based and, iii) clustering-based detection.

Given the diversity of theft scenarios and associated attack vectors over different data aggregation levels on the smart grid infrastructure, detection methods have been employed either at a centralised or a distributed fashion. Table 2.5, Table 2.6 and Table 2.7 provide a comprehensive summary of methods introduced in past literature over the last decade. Evidently, the majority of methods consider a combinatorial use of algorithmic techniques in order to address specific challenges ranging from data pre-processing and filtering up to statistical correlation analysis. Furthermore, some formulations are broadly used (e.g., artificial neural networks -ANNs and support vector machines - SVMs) over different types of attacks operating under diverse data types gathered at various smart grid data aggregation components.

Complementary, Table 2.8, Table 2.9 and Table 2.10 illustrate the experimental approach underpinning the methods summarised in Table 2.5, Table 2.6 and Table 2.7 further provide their outcomes. As depicted, each method was employed over energy theft use cases involving a number of nodes within the actual grid and utilised specific statistical features. In summary, we identify a range of raw as well as post-processing features that are utilised within the listed methods. Thus, there exist techniques involving one or more of basic statistical features (e.g., mean, min/max), frequency and temporal domain features (e.g., signal periodicity frequency components), scaling on independently distributed raw data, clustering or probability-based similarity metrics as well as locality (e.g, geolocation coordinates), auxiliary (e.g., number of energy appliances) and environmental features (e.g., temperature, humidity).

2.3.2.1 Classification-based detection

The study of (Messinis, Rigas, and Hatziargyriou, 2019b) proposed a classification system to detect energy theft conducted at the end user infrastructure. The introduced solution was assessed over simulations replaying the Irish Smart Energy Trail dataset and its operation relied on the synergistic use of an SVM classifier, a power optimization scheme and a voltage sensitivity analysis component. However, this system required the utilization of additional features such as voltage and active energy data to detect theft. The problem with utilizing such sensitive measures is that it can expose customer data to privacy violations. Moreover, features associated to real-time ancillary services (e.g., active/reactive energy require adequate signal smoothing techniques for complete conversion over the time-frequency domain; an element missing from this piece of work as it is not encapsulated within SVM formulations or the proposed pre-processing stage.

Table 2.5: Overview of the data-driven energy theft detection methods.

Ref.	Technique	Nature		Attack Infrastructure	Attack Type	Data Type
		Centred	Distributed			
(Fernandes et al., 2018)	OPF, SVM, Bayesian Classifier, Logistic Regression	√		T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(Y. Peng et al., 2021)	Local Outlier Factor, K-means, Maximal Information Coefficient, Clustering by Fast Search and Find of Density Peaks	√		End user	Demand data manipulation	Consumption
(Messinis, Rigas, and Hatziargyriou, 2019b)	SVM, Voltage Sensitivity Analysis, Breakout Detection Package	√		End user	Demand data manipulation	Consumption
(Meira et al., 2017)	Random Forest, Logistic Regression, SVM, K-means	√		T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(Glauner, Meira, Dolberg, et al., 2017)	SVM, K-NN, Random Forest, Logistic Regression	√		T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(Aydin and Gungor, 2018)	Logistic Regression, K-NN, Fourier Transform, Random Forest		√	T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(Gunturi and Sarkar, 2021)	CatGBM, Random Forest, AdaBoost, LightGBM, Extra Trees, XGBoost, Near-miss,SMOTE,	√		end user	Demand data manipulation	Consumption
(Z. Zheng et al., 2018)	Wide & Deep CNN, Three-Sigma Rule, Random Forest, CNN, SVM, Logistic Regression	√		End user	Demand data manipulation	Consumption
(Ying Zhang, J. Wang, and B. Chen, 2020)	Autoencoders, Generative Adversarial Networks, SVM, K-NN		√	T&D	SCADA data manipulation	Network measurements

Table 2.6: Overview of the data-driven energy theft detection methods (Con.).

Ref.	Technique	Nature		Attack Infrastructure	Attack Type	Data Type
		Centred	Distributed			
(Yao et al., 2019)	Convolutional ANN, Paillier Algorithm, SVM, Random Forest, Logistic Regression	√		End user	Demand data manipulation	Consumption
(Ashrafuzzaman, Das, et al., 2020)	Logistic Regression, DT, ANN,SVM, Naive Baye, LOF, Isolation Fores, Elliptic Envelope	√		T&D	SCADA data manipulation	Network measurements
(Punmiya and Choe, 2019)	XGBoost, CatGBM, LightGBM	√		End user	Demand data manipulation	Consumption
(Shaaban et al., 2021)	Linear Regression, SVM, DT		√	Generation	Generation data manipulation	PV measurements
(W. Li et al., 2019)	MLPNN, RNN, LSTM, GRU, Simple Moving Average		√	T&D and end user	Demand data manipulation (Direct tapping)	Home appliances data
(Gao, Foggo, and Yu, 2019b)	Linear regression, SVR, ANN, Radial Basis Function Network	√		End user	Demand data manipulation	Consumption
(Razavi, Gharipour, et al., 2019)	Finite Mixture Clustering, Genetic Programming, ANN, Random Forest, SVM, K-NN, GBM	√		End user	Demand data manipulation	Consumption
(Buzau et al., 2018)	XGBoost, K-means, K-NN, SVM, Logistic Regression	√		Generation and T&D and end user	Attacks caused NTL	Consumption
(Hegazy et al., 2022)	Temporal CNN, LSTM, CNN		√	T&D	SCADA data manipulation	Network measurements
(K. Zheng et al., 2018)	Maximum Information Coefficient, CFSFDP, Pearson Correlation Coefficient, Kraskov's Estimator, LOF, FCM		√	End user	Demand data manipulation	Consumption
(Ashrafuzzaman, Chakhchoukh, et al., 2018)	Deep Learning, Generalized Linear Modeling, Random Forest, GBM	√		T&D	SCADA data manipulation	Network measurements

Table 2.7: Overview of the data-driven energy theft detection methods (Con.).

Ref.	Technique	Nature		Attack Infrastructure	Attack Type	Data Type
		Centred	Distributed			
(Esmalifalak et al., 2017)	SVM, Density based anomaly detection, PCA		√	T&D	SCADA data manipulation	Network measurements
(Mahmoud et al., 2020)	Deep Feed Forward ANN, Deep Recurrent ANN, Deep Convolutional Recurrent ANN, SVM, ARIMA		√	Generation	Generation data manipulation	PV measurements
(Nallathambi, 2017)	Random Forest, DT		√	T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(M. Wen et al., 2021)	CNN	√		End user	Demand data manipulation	Consumption
(Mukherjee, Chakraborty, and Ghosh, 2022)	CNN, SVM, LightGBM, ANN		√	T&D	SCADA data manipulation	Network measurements
(Jindal, Dua, et al., 2016)	DT, SVM		√	T&D and end user	Attacks caused NTL	Consumption
(Cody, Ford, and Siraj, 2015)	DT	√		T&D and end user	Demand data manipulation (Direct tapping)	Consumption
(Jokar, Arianpoo, and Leung, 2016)	SVM, K-means	√		End user	Demand data manipulation	Consumption

Table 2.8: Experimental approaches of surveyed studies on data-driven energy theft detection.

Ref.	Number of Nodes (≈)	Features	Evaluation Metrics (Best algorithm) (%)	Experimental Evaluation		Percent of Attacked Samples (≈) (%)
				Simulation	Testbed	
(Gunturi and Sarkar, 2021)	—	Statistical	AUC = 90,PR = 99,RE = 98,F1 = 75		√	$10-50$
(Glauner, Meira, Dolberg, et al., 2017)	$700k$	Locality, Auxiliary	AUC = 62.8		√	$1-90$
(Jokar, Arianpoo, and Leung, 2016)	$5K$	Auxiliary, Similarity	FPR = 0.1, TPR = 94		√	–
(Punmiya and Choe, 2019)	$5k$	Statistical	FPR = 4, TPR = 97		√	50

Table 2.9: Experimental approaches of surveyed studies on data-driven energy theft detection (Con.).

Ref.	Number of Nodes (≈)	Features	Evaluation Metrics (Best algorithm) (%)	Experimental Evaluation		Percent of Attacked Samples (≈) (%)
				Simulation	Testbed	
(Messinis, Rigas, and Hatziargyriou, 2019b)	$5K$	Statistical, Auxiliary, Scaling, Frequency	ACC = 99.4, FPR = 0 TPR = 98.9, AUC = 99.9	√	√	50
(Yao et al., 2019)	$42k$	Similarity	ACC = 92.67		√	—
(Z. Zheng et al., 2018)	$42k$	Statistical, Scaling	AUC = 96.86		√	9
(Nallathambi, 2017)	1	Auxiliary, Environmental, Temporal	ACC = 95.78, AUC = 100		√	–
(Shaaban et al., 2021)	400	Environmental, Scaling	ACC = 91.50, FPR = 11.5, PRE= 89.15		√	–
(Razavi, Gharipour, et al., 2019)	$4k$	Statistical, Similarity	ACC = 99, AUC = 99.8		√	–
(Y. Peng et al., 2021)	$3.5k$	Statistical, Similarity	AUC = 91.84		√	12
(K. Zheng et al., 2018)	391	Statistical, Auxiliary	AUC = 81.6		√	12.8
(Mukherjee, Chakraborty, and Ghosh, 2022)	180	Statistical, Scaling	ACC = 97, PRE = 99.53, RE = 99.79, F1 = 99.64	√		9
(Hegazy et al., 2022)	180	Statistical, Scaling	PRE = 99.83, RE = 99.92, F1 = 99.87	√		—
(M. Wen et al., 2021)	$42k$	Statistical, Scaling	ACC = 91.9, AUC=79.1	√		—
(Ying Zhang, J. Wang, and B. Chen, 2020)	119	Statistical, Scaling	ACC = 97.85, PRE = 92.68, RE = 90.49	√		50
(Ashrafuzzaman, Chakhchoukh, et al., 2018)	$100k$	Auxiliary	ACC = 97.7, F-score = 98.78 AUC = 98.53	√		—
(Mahmoud et al., 2020)	71	Auxiliary	TPR = 99.3, FPR = 0.22 F-score = 99.55	√		—
(Cody, Ford, and Siraj, 2015)	$5k$	Temporal	–		√	–
(Meira et al., 2017)	$3.5M$	Similarity, Temporal, Locality, Auxiliary	AUC = 75.03		√	$10 - 90$
(Aydin and Gungor, 2018)	425	Statistical, Frequency, Scaling	ACC = 98.37, FPR = 0 F-score = 87.50		√	16
(Buzau et al., 2018)	$57k$	Statistical, Similarity, Auxiliary	AUC = 91		√	$5.38 - 8.37$

Table 2.10: Experimental approaches of surveyed studies on data-driven energy theft detection (Con.).

Ref.	Number of Nodes (≈)	Features	Evaluation Metrics (Best algorithm) (%)	Experimental Evaluation		Percent of Attacked Samples (≈) (%)
				Simulation	Testbed	
(Esmalifalak et al., 2017)	$1k$	Similarity, Auxiliary	F-score = 95	√		—
(Ashrafuzzaman, Das, et al., 2020)	$100k$	Statistical	F1=84, ACC =89,PR=99, FPR= 0.03, AUC=86, TPR=73	√		–
(W. Li et al., 2019)	1	Statistical	ACC = 99.96		√	—
(Gao, Foggo, and Yu, 2019b)	980	Auxiliary	–		√	–
(Fernandes et al., 2018)	$42k$	Statistical, Auxiliary	ACC = 83, F-score = 80.9		√	—
(Jindal, Dua, et al., 2016)	$1k$	Scaling, Auxiliary, Environmental, Temporal	ACC = 92.5, FPR = 5.12	√		20

Variations of the conventional SVM formulation in synergy with principal component analysis (PCA) was also the basis behind the work (Esmalifalak et al., 2017). The evaluation of SVM-based formulations was based on labelling load data that were simulated as stochastic processes such as to comply with pragmatic power system behaviour in the T&D system infrastructure. PCA was initially employed in order to reduce the high dimensionality of the simulated measurements and they were firstly labelled within the training process of a supervised SVM formulation. Subsequently, newly generated measurements were tested over the supervised model and the identification of outliers implying theft detection was feasible with 95% accuracy. However, due to the dependence of the proposed scheme on PCA, there exists a high likelihood of a trade-off between the loss of important information included in the simulated measurements and the dimensionality reduction process.

Recent developments in the area of deep learning enabled the composition of adequate energy theft detection schemes. In (Yao et al., 2019) a novel synergy of convolutional neural networks (CNN) and the Paillier cryptosystem in order to maintain user privacy but also detect energy theft was demonstrated. Under a similar mindset, a modified wide and deep CNN was proposed in (Z. Zheng et al., 2018) in which the wide component of the customised CNN deals with global consumption features whereas the deep CNN component was more focused on profiling the consumer's consumption periodicity such as to detect deviations implying energy theft at end user level. In (M. Wen et al., 2021), a novel privacy-preserving energy theft-detection framework utilising federated learning-based CNN was introduced to

detect theft in consumption measurements. Nonetheless, the use of federated learning requires local detection points and the addition of new components, or customers in the grid would negatively affect both scalability and accuracy performance.

Nevertheless, a deep learning-based approach was also proposed to detect theft activities in T&D infrastructure in (Ying Zhang, J. Wang, and B. Chen, 2020). The proposed method utilizes autoencoders to efficiently reduce the data set's dimensions and extract features. It also incorporates autoencoders into generative adversarial networks, which construct an adversarial game between two ANNs. The proposed approach effectively achieves a detection accuracy of greater than 94%. In (Mukherjee, Chakraborty, and Ghosh, 2022) a CNN-based multi-category classifier was proposed to capture discrepancies in energy-flow measurements caused by possible theft attacks. The method proposed in this study was successful in achieving a high detection accuracy of 97%. A similar combination was adopted in (Hegazy et al., 2022), where a parallel approach based on the LSTM with temporal CNN is proposed. With an F1 score of 99.87%, exhaustive simulations indicate that the suggested method can detect the presence and location of theft attacks in T&D systems.

The superiority of deep learning-based energy theft detectors was also illustrated at the work in (Ashrafuzzaman, Chakhchoukh, et al., 2018) where a number of traditional and ensemble classifiers such as random forests, and gradient boosting machines (GBM) were compared with a CNN-based classifier using T&D infrastructure measurements. Similarly in (Mahmoud et al., 2020) the applicability of a deep learning-based detection solution based on measurements that are captured at DRES deployments was demonstrates. However, such theft detection methods entail enormous computational costs due to the large amount of data required to effectively train fully supervised deep learning-based detectors.

Several studies have also provided insightful comparisons of various classification-based energy theft detection schemes and insights on the performance of particular statistical features. For instance, the work in (Fernandes et al., 2018) introduces the use of a customised optimum path forest (OPF)-based detection scheme for attacks that target explicitly energy theft. In evaluations of industrial and end user consumption data the proposed scheme outperformed conventional classifiers such as SVM and Bayesian classifiers with respect to detection accuracy. However, with respect to log loss function, SVM achieved the best value, outperforming the customised OPF-based detection scheme. In (Meira et al., 2017), examine a diverse set of spatiotemporal and exogenous features based on four criteria, namely, auxiliary, similarity, locality and temporal. The performance of the selected features was investigated through the classification processes of customised SVM, logistic regression and random forest formulations. It was clearly revealed that features derived only from consumption measurements (such as similarity features) are adequate for the accurate detection of energy theft attacks. However, such a

detection study entails computational processes on further features from historical consumption measurements, which limits the application of this method in large-scale detection scenarios.

In parallel, the study in (Glauner, Meira, Dolberg, et al., 2017) demonstrates that the classification process under various algorithms (e.g., SVMs) reveals that features related to aggregated neighbourhood consumption alongside locality parameters outperformed individual meter time series distributions. However, we argue that energy theft detection based on the utilization of features related to neighbourhood consumption and locality parameters may not be generic enough, due to the fact that the consumption patterns of those who belong to the same geographical domain differ from one another. The assessment of features pointing to energy theft in synergy with classification performance were also one of the main focus areas in the studies conducted in (Aydin and Gungor, 2018; Buzau et al., 2018; Gunturi and Sarkar, 2021) and (Punmiya and Choe, 2019). These analyses demonstrated that combining a detection classifier with feature engineering not only enhanced detection performance but also decreased data sorage space and processing time for the detection process.

Through the application and comparison of classification-based ensemble methods (e.g., XGBoost, CatGBM, LightGBM) with conventional classifiers (e.g., ANNs, SVMs) over simulated attack scenarios it was revealed that ensemble methods contribute significantly towards computationally-efficient and more accurate theft detection (Punmiya and Choe, 2019). An ensemble learning-based approach was also proposed to detect theft attacks on T&D infrastructures in (Ashrafuzzaman, Das, et al., 2020). In this study, classification-based models (e.g. SVM, ANN, decision tree (DT)) are used in one ensemble, while anomaly-detection models are used in the other (e.g. LOF and one-class SVM). In each scenario, the ensemble scheme's results were compared to those of corresponding individual models. The performance of a single classification-based model is equivalent to that of ensemble models. For models based on anomaly detection, however, ensemble performance was superior to that of individual models. However, ensemble-based detection methods pose some instability since a slight variation in the training data would unavoidably entail substantial restructuring of the main tree-based detection model. Thus, imposing higher computational costs. Nonetheless, the work in (Ashrafuzzaman, Chakhchoukh, et al., 2018) demonstrates the superiority of deep learning-based theft detection schemes over any ensemble-based approaches compared, where the detection accuracy based on the deep learning technique was 97.7%.

Despite the relatively high accuracy performance and reliability of classification-based techniques, the aforementioned detection methods require labelled data from malicious and energy theft-free behaviours. Obtaining such data is either challenging in a real scenario or, even if they exist, they do not cover all possible theft-attack behaviours (Messinis and Hatziargyriou, 2018). Theft-free data can be collected from

historical grid measurements, however, malicious data (i.e., theft samples) covering the spectrum of theft behaviours for a particular node hardly exist. In such cases, the performance of the detection method is limited due to malicious sample unavailability. These methods may remain unsuccessful in detecting more advanced and stealthy attacks that are not available in training data, which directly affects the overall detection performance (Júnior et al., 2016).

2.3.2.2 Regression-based detection

The study in (W. Li et al., 2019) proposed a modular energy theft detection system consisting of a three-stage decision making process achieving 99.96% on theft detection accuracy. The first stage relies on a multi-model power consumption prediction system based on multi layer perceptron neural network (MLPNN), Long short term memory (LSTM) ANN, recurrent neural network (RNN) and gated recurrent unit (GRU). The second stage deals with monitoring a moving average whereas the third stage employs a customer's historical measurements to determine occasional maximum energy consumption in order to make a final decision on a theft attack. Although interesting results are achieved, the proposed method is undynamic for any future changes in consumption patterns, since the main focus of such a system is the utilization of historical consumption measurements in the detection process.

The behavioral profile of normal energy consumption was assessed in (Cody, Ford, and Siraj, 2015) in order to detect deviations implying energy theft. The conducted experiments revealed that consumption values can be predicted using DT learning and they can be categorised into normal or fraudulent based on the threshold root mean squared error value. Any value exceeding this threshold indicates a possible energy theft attack. However, the prediction formulation proposed in this study can be improved through the utilization of further comprehensive features, such as numbers of appliances and providing the prediction model with additional details to determine consumers' energy consumption patterns.

A synergistic use of SVMs and DT for theft detection in end user infrastructure was proposed in (Jindal, Dua, et al., 2016). Decision-tree formulation operates on various features, including the numbers of heavy appliances and persons, to generate the predicted consumption of each consumer. Then, an SVM-based classifier is used to detect malicious consumers. Results show that the proposed method can be implemented in real-time scenarios as the false positive rate is significantly reduced to 5.12%. Complementary work in (Nallathambi, 2017) achieves regression based on random forests to predict the expected energy consumption over the US-wide consumption profiles for 2014. Through the use of various performance metrics (e.g., prediction accuracy, classification error rate) forecasting through random forests achieved 95.78% of prediction accuracy and outperformed a DT-based

approach that reached 91.6% accuracy. Thus, providing a quite effective energy theft prediction scheme. However, such a scheme cannot be considered as generic since energy consumption is usually characterized by invariable variance or non-stationary behaviour. Therefore, the fundamental principles underpinning random forests model could become inappropriate for identifying short-term irregularities in energy consumption.

Nevertheless, an anomaly detection-based approach entailing a regression tree model and probability density function is developed in (Shaaban et al., 2021). In the training phase of the detector, historical records of solar irradiance, temperature, and smart meter readings are employed. As a metric for detecting suspicious data, the probability density function of the difference between the actual readings from DRES meters and the anticipated generation by the regression model is applied.

A data-driven regression model was proposed in (Gao, Foggo, and Yu, 2019b) for energy theft detection. Instead of using unreliable topology information and parameters from secondary network, this method was based on modified linear regression algorithm. It uses only the voltage data and consumer's consumption data making it more feasible to adopt. Finally, the training data from real world smart meter was used to validate proposed method and results illustrate effective identification of cases related to energy theft. However, customers' data may be vulnerable to privacy breaches due to the dependence upon voltage measurements.

Overall, despite the applicability of the aforementioned methods to identify advanced energy-theft attacks, regression-based methods regularly demonstrate longer detection times than other detection categories. In such cases, regression techniques are principally employed in the first stage of theft-detection methods and require additional procedures to reach a final decision during the detection process. This in turn is a time-consuming task and limits the applicability of such methods in a real-time energy trading scenario, where the time required to detect theft activities is influential in preventing any losses.

2.3.2.3 Clustering-based detection

A clustering-based theft detector utilising consumption patterns was also proposed in (Jokar, Arianpoo, and Leung, 2016). In order to improve classification accuracy, the number of clusters in the examined dataset was filtered through Silhouette plots and subsequently clusters were hierarchically labelled across various consumption profiles. The resulted outcomes of this approach demonstrate that even with low measurement sampling intervals, the algorithm is scalable and achieves a detection rate of 94%. However, the proposed technique required the installation of transformer meters, which increased the monetary cost of such systems.

An alternative approach based on genetic algorithms and finite mixture modeling

for composing clusters of consumption in order to identify customer segmentation and potential outliers was presented in (Razavi, Gharipour, et al., 2019). In fact, the proposed method outperforms a number of classification-based approaches such as k-nearest neighbours (K-NN), ANN and SVM by 99.8% in the area under the curve for theft detection. However, such a detection system cannot be applied in a real-time scenario, since the results achieved indicate that there is an increase in the relative to physical inspection.

An outlier-based detector of three modules was presented in (Y. Peng et al., 2021). The proposed method applied local outlier factor (LOF) and the K-means algorithm as the basis to detect theft at the end user infrastructure. Firstly, consumption profiles were analysed with k-means and subsequently outlier candidates were selected based on the deviation of each consumer from the relative cluster centers. Finally, the anomaly ranking of the selected candidates was calculated using the LOF algorithm. Although the proposed detector achieves reasonably high detection accuracy of 91.84%, it still fails to detect linear theft, where an attacker manipulates the consumption profile to reduce it at a constant rate.

In addition to consumption measurements, the authors in (K. Zheng et al., 2018) employed the measurements recorded by an observer smart meter, installed to aggregate the sum of the consumption measurements of a group of consumers over a certain period. The proposed approach in (K. Zheng et al., 2018) combined two data-driven techniques, i.e. a maximum-information coefficient and a clustering technique by fast search and finding density peaks (CFSFDP), to detect these thefts.

Despite the fact that clustering-based methods can be used in scenarios of scarcity, minimal or zero availability of malicious intent, these methods will normally produce an end result with a high false-positive rate. To construct a clustering-based model, no assumptions of labelled data from malicious and theft-free behaviours are made. As a result, the detection model can identify any abnormal patterns as malicious behaviours (Messinis and Hatziargyriou, 2018). In general, abnormalities may occur due to non-malicious activities (e.g., smart-meter misconfigurations), leading to an increase of false-positive rates resulted by clustering-based theft detection mechanisms.

2.3.2.4 Comprehensive analysis

Undoubtedly, malicious actors continue to target a diverse set of vulnerabilities present over various system, network and algorithmic components serving the (sub)infrastructures composing a smart grid deployment. Hence, attackers intend to launch energy theft attacks through a variety of techniques that target the evasion from current detection schemes. Evidently, data-driven methods for detecting energy theft distilled by learning, profiling and detecting abnormalities are considered as a means to adaptively engage with new attack vectors.

In general, data-driven energy theft detection schemes leverage three conceptual and data-driven procedures; (i) data-processing and model-selection stages covering aspects of data sanitisation and feature selection, (ii) model-training procedure which varies across classification, clustering and regression detection methods and (iii) decision-making procedure which includes applying a model trained on new data such as to pinpoint anomalies that could relate with malicious activity.

Given the "ad-hoc" employment of most of the detection methods presented herein over specific use cases, we argue that there is no universal data-driven methodology covering all aggregation levels in a given smart grid deployment. In general, the aforementioned three levels, categorized into energy generation, T&D, and end user infrastructures have different probabilities for the deployment of theft and different vulnerabilities exploited by malicious actors. Such factors should be taken into consideration when a method is designed to detect energy-theft attacks.

However, the utilization of a hybrid data-driven model has proven to be more robust than adopting a single model in detecting attack vectors underpinning energy theft. Such hybrid methods are considered to make combinatorial use of two or more data-driven models. In such methods, the entire theft detection method leverages the analytic process of each candidate model to achieve a specific action. All achieved actions are subsequently integrated into one detection system in order to complement each other and mitigate the limitations of the others.

Furthermore, the utilization of data from multiple and diverse sources can create a more reliable method for detecting energy-theft attacks over smart-grid infrastructures. Detection methods utilizing a single data source are constrained to build a candidate model fitting specific data measurements, thus its suitability is not generic. Moreover, the candidate model is sensitive to the samples it was trained with, which may potentially have been manipulated to falsify the detection method to cope with new adversarial objectives. However, by acquiring the data from various sources which have less likelihood to be accessible to adversaries can significantly increase the reliability and performance of the detection method.

The adoption of data-driven methods that utilise multiple and diverse data feeds would unavoidably invoke trade-offs spanning across performance, privacy and computational complexity. For instance, data-driven theft detection at the end user infrastructure method would require a privacy-aware data processing and aggregation scheme. Hence, in order to detect theft in DRES infrastructure, the detection method should not rely on data that are not available to utility providers such as EMS measurements. Such measurements are usually maintained by the DRES owner and not accessible to any third party. Thus, there could be some limitations in terms of the granularity of the anomaly detection process employed by the theft detection scheme. On the other hand, energy theft-detection process in the T&D infrastructure inherently requires the utilization of high volume of network and system

log measurements. Therefore, an anticipated high computational cost would be implied and thus limit the real-time capabilities of a given theft detection scheme.

2.4 Present challenges and suggested solutions

Despite the various solutions proposed in terms of energy theft detection, there exist various gaps and open issues thus requiring further attention within suggested solutions. Within this section, we highlight and discuss some of the challenges and we further summarize potential suggested solutions. As depicted in Fig. 2.6 we decompose the gaps spectrum into (i) measurement-driven, (ii) machine learning and (iii) security-related challenges.

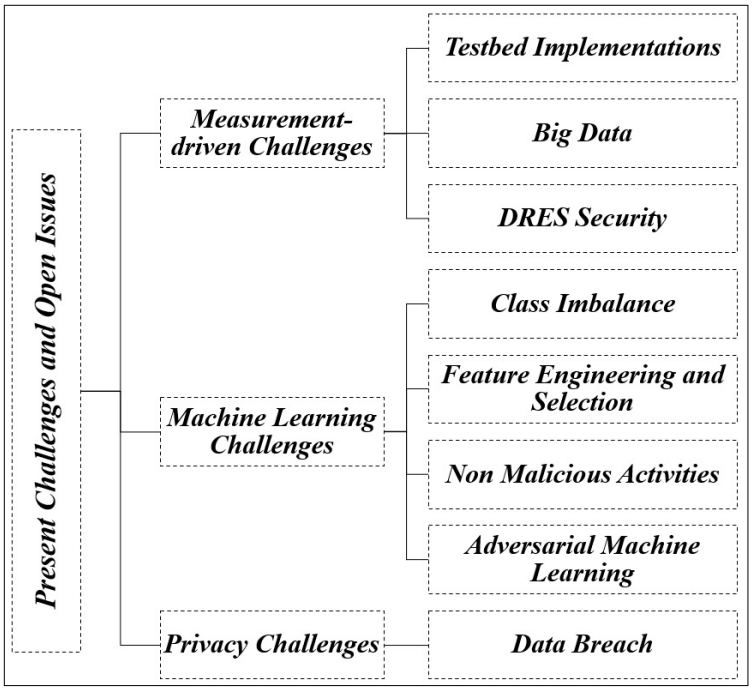

Figure 2.6: Present gaps in energy theft detection.

2.4.1 Measurement-driven challenges

2.4.1.1 Testbed scenarios and datasets

Diverse energy-related data sets, different network infrastructures, and multi-faceted energy theft-related attacks are studied in most of the presented works as discussed in this study. However, there is a notable lack of commonly available (and applied)

prototype implementation on realistic large-scale testbed as well as datasets such as to enable extensive experimental verification nor experimental reproducibility tailored for energy theft detection (Messinis and Hatziargyriou, 2018). Most testbeds and their corresponding datasets are principally designed in an ad-hoc fashion for specific projects limiting the generalisation of findings (Cintuglu et al., 2016). Therefore, we argue that it is of crucial importance to build benchmark testbeds and properly designed platforms such as to test connections and security features of a system and maintain alignment with the pragmatic and rapidly emerging design requirements of current and future smart grid deployments.

2.4.1.2 Measurements and big data

Volume, velocity, and variety are the traditional traits and they naturally challenge any analysis domain within the smartgrid ecosystem (J. Hu and Vasilakos, 2016). Hence, the adequate comprehension and optimisation of these diverse traits during data collection, processing and analysis over particular smartgrid scenarios such as energy theft detection is of vital importance. For instance, there are 27 million consumers that consume domestic electricity in the United Kingdom alone. These consumers have more than 100 million data points that are collected either quarterly or half-yearly. These points are used by the energy suppliers to store, record and use in the billing system and identifying abnormal conditions that could relate to specific energy theft-related attacks. However, with smart metering, to collect the data from these many data points, at a thirty-minute sampling rate, will require a substantial amount of resources. For example, at least 4500 to 9000 times more of the present data size will be required to be processed by the energy suppliers, and therefore this leads to a significant augmentation in data size (Wilcox et al., 2019). Thus, there is a strong requirement for efficiently coupling the measurement requirements for granular energy monitoring with optimised storage as well as data processing solutions.

2.4.2 Machine Learning challenges

2.4.2.1 Class imbalance

Class imbalance problem is a traditional problem existing for supervised or semi-supervised learning having direct implications on energy theft detection. In particular, this problem occurs when one of the classes (in a multi-class problem) has significantly more number of samples than the other classes, thus the training model is biased leading the testing phase to classify events towards the majority class label (Gunturi and Sarkar, 2021). Hence, in the case of learning for theft instances in which are by far less than legitimate instances, the class imbalance problem would result on a classifier to incorrectly label malicious instances to the majority of normal behaviour. It is

therefore important to establish adequate ground truth datasets with correct scaling factors through the training phase of learning processes by assigning correct weight parameters to malicious samples. Thus, addressing the limitations from the class imbalance problem (Maamar and Benahmed, 2018). Nonetheless, the composition of concrete ground truth labels for theft instances is also a topic aligned with the needs of optimised feature engineering and selection as we discuss next.

2.4.2.2 Feature engineering and selection

Feature engineering accompanied by efficient feature selection is a powerful foundation for addressing the aforementioned class imbalance problem as well as tailoring a learning procedure to identify energy theft instances. Evidently, it is common in many energy theft detection processes to operate over insufficient or incomplete feature vectors and experience class imbalance as well as model over-fitting (i.e., learn the only specific pattern in a given dataset), thus affecting significantly detection accuracy. Therefore, designing and engineering new features can improve the performance of machine learning detection methods (Glauner, Meira, Valtchev, et al., 2016; Maamar and Benahmed, 2018).

2.4.2.3 Non-malicious abnormal activities

A classical problem within anomaly detection is the distinction of classes between anomalous events. Energy theft-related attacks could relate to statistical abnormalities and have similar properties as anomalous events that are caused by legitimate intent (e.g., smart meter misconfiguration). A great challenge is to compose adequate classification and clustering schemes that are able to pinpoint the differences between malicious and legitimate processes and further highlight the specific properties entailed within an energy theft incident. There can be many reasons that the ambiguities in electrical node output patterns may occur. These can happen owing to several altered causes such as new device installation (for example, a new DRES) or changing in the electricity usage habit of the residential end users (Messinis and Hatziargyriou, 2018). This, in turn, increases the overall inspection cost (Jindal, Dua, et al., 2016) as once the model classifies an energy theft attack, physical inspection is essential for final verification and that is a costly procedure (Jokar, Arianpoo, and Leung, 2016). It can, therefore, be argued that there is a requirement for more research in the improvement of the proposed detection methods in terms of reconsigning the theft detection activities and reducing the false positive alarms (Maamar and Benahmed, 2018).

2.4.2.4 Adversarial machine learning

As already described, it is feasible for an adversary to manipulate end user data or game the algorithmic learning procedure in a targeted manner. These particular types of attacks are called adversarial machine learning attacks which are carried out for the purpose of theft detection. For example, carrying out an attack where input data is made to look like normal electrical data, i.e., crafting an attack that seems normal to the machine learning algorithm or changing the weights of the trained machine learning model (Jokar, Arianpoo, and Leung, 2016). These scenarios can maximize the predicted loss or falsify trained models to new adversarial objectives (Bor et al., 2019; L. Huang et al., 2011; Y. Chen, Y. Tan, and Deka, 2018). Moreover, handcrafted rule-based attacks are more sophisticated (than automated attacks) and proposes different challenges, and therefore a generalized detection model will not provide promising results (Bor et al., 2019). Thus, more studies are required to investigate the capabilities and the limitations of existing machine learning detection algorithms with respect to adversarial machine learning.

2.4.3 Privacy challenges

2.4.3.1 Data breach

Most of the energy theft attack detection schemes utilize (some of) the private information of consumers/prosumers, such as smart meter readings and user load/generation profiles. While this information can help to detect the theft attacks to a certain extent, it should still be kept in mind that disclosing such private data may raise concerns about the user's safety and breach his/her privacy. These data breach threats can occur in different stages of the theft detection process, including data collection, transmission and storage. Such sensitive breached information might be purchased by interested third parties such as marketing companies which can use this data to sell their products to possible customers. Apart from this, if criminals get their hands on this sensitive data, the daily routine of a household can be analyzed from electricity usage/generation pattern to carry out crimes. Therefore, detecting energy theft attacks while maintaining privacy of information is a challenging task, but there is a notable lacking of intelligent privacy-preserving detection schemes in the works of the energy theft (S. Salinas, Ming Li, and P. Li, 2012).

2.4.4 Measurement-driven solutions

2.4.4.1 Testbed simulation, emulation and hardware

Future works should consider the measurement-driven challenges that affect energy theft detection frameworks. The energy theft activities should be ratified by experimental environments and for this to happen, there is a strong need to include testbed software simulation, emulation and hardware for carrying out energy theft analysis. For instance, a cloud-based environment can be created to store smart grid data which can be used in these testbeds to conduct energy theft analysis (S. Tan et al., 2017). With simulation software and emulation hardware, a quick verification of new concepts can be achieved efficiently which can then be easily transferred to power system industry and for more extensive public use. Moreover, these testbeds create interesting educational platforms to understudies which would spur the research interests to conduct multi-user experimental facilities for several smart grid applications (Cintuglu et al., 2016).

2.4.4.2 Big data schemes

To collect, store, and process monitoring data various diverse data sources in smart grid results to the big data challenges as discussed earlier. To cater to these challenges the two important present challenges include the creation on big data analysis platforms and reducing the complexity of such data. For the former, cloud computing technology has been used to create big data platforms by the many industries since this technology is scalable, self-organizing, and adaptive. Therefore, platforms such as Hadoop, Cassandra, and Hive in conjunction with cloud computing can be used by utility providers for smart grid big data analysis (Bhattarai et al., 2019). For the latter (to reduce the data complexity), different techniques such as dimensionality reduction, distributed optimization algorithms, and active learning can be useful to analyze big data efficiently (Meng Li et al., 2020). Different studies reported that the computational process of the summarized and produced data rather than the original data stream can result in an acceptable relative error (Jindal, Kumar, and M. Singh, 2020a). Therefore, these dimensionality reduction techniques are useful for reducing the communication cost, computing complexity, and storage resource utilization for smart grid big data analysis (Diamantoulakis, Kapinas, and Karagiannidis, 2015).

2.4.5 Machine learning solutions

2.4.5.1 Class imbalance

Class imbalances happen when there are less samples in one of the target classes for machine learning algorithms or a close similarity in the number of samples in

considered classes. To enhance the learning results associated with imbalanced data classes (and improve on their bias), three primary methods can be utilized: data-level, algorithm-level and hybrid techniques (Krawczyk, 2016). In the data-level techniques, the concentration is on the modification of training set to allow more balanced distributions for oversampling (more minority groups' samples) and undersampling (fewer majority groups' samples). The algorithm-level techniques modify the learners that already exist to eliminate their bias for majority groups. However, good insight is required into the modified learning algorithm and real discovery of reason for skewed mining distributions. Some popular algorithmic techniques include cost-sensitive approach (to insert different penalties for every group of samples) and one-class learning (concentrating on the specific target groups). The hybrid techniques use the combination of methods as mentioned above, by reducing their weaknesses and making use of their strengths (Krawczyk, 2016).

2.4.5.2 Feature engineering and selection schemes

We argue that future research directions could place stronger focus on particularly exploring algorithmic and system-wide principles to facilitate automated feature engineering and selection methods. The feature engineering process can extend the original detection model's feature vector by adding new features that are calculated based on other input features. These engineered features may be the differences, averages, or other statistical transformations of the original feature vector, helping in better understanding of the interactions amongst these features. This process is similar to the statistical transformations performed by human analysts for constructing an engineered feature formulas. The task of feature engineering and selection is mainly a time-consuming task and each model type will respond in different manner to different engineered feature types (Heaton, 2016). However, in general, the selected and engineered featured would help in achieving the maximum probability of success for the machine learning algorithms to detect energy theft (Jundong Li et al., 2018). Typically for feature engineering and selection, many methods can be used such as mathematical functions, deep feature synthesis components, expansion reduction, evolution-centric, multi layer neural networks and hyper parameter optimization (Heaton, 2016; Khurana, Samulowitz, and Turaga, 2018).

2.4.5.3 False positive rate-reduction schemes

A meta-learning scheme can be helpful to reduce the false positive rates resulting from non-malicious activities in the process of energy-theft attack detection. Meta-learning can be defined as a learning process involving the collection of knowledge from past experience in order to use it in future learning (Jinghang Li and M. Hu, 2020). Meta-learning is required by the theft-attack detection system to combine various

classifiers (by taking note of their behaviours) and adopting an integration rule to reduce false positives. In the literature, the main meta-learning techniques include stacking, bagging, voting and boosting. In the voting approach, each classifier has one vote, and the classification that has the highest votes determines the final prediction. In stacking learning, the process adopts a layered architecture wherein each layer has one or more classification techniques. A layer's projection is applied to extend the original vector of the feature with the closest instance. The bagging approach creates a combination of classifiers through the manipulation of training samples in a base classifier. It selects one base classifier and invokes it many times using several training samples. Boosting, in contrast to bagging learning, generates various basic classifiers through a procedure in which examples of data sets receive new weights in sequence (Possebon et al., 2019).

2.4.5.4 Adversarial machine learning schemes

With respect to adversarial machine learning, a binary classifier-based intrusion detection system trained on available device behaviour logs is imperative (Bor et al., 2019). This system can attempt to tag approaching instances as either malicious or benign, using features which are generated in real-time from streams of energy data. Through gradual training instances expansion and feature generation refinement, this system can produce a confidence score that can be utilized to set recall/precision. This will allow having low maintenance overheads and fewer false alerts as compared to a manual system. The underlying intrusion detection system can employ a broader range of features including outgoing data from the control algorithm (Bor et al., 2019). As also discussed in (Bor et al., 2019; Jindal, Angelos K. Marnerides, et al., 2019) malicious behaviours can be detected using other associated features such as network properties (e.g. packet size, packet arrival time) and communication security (e.g. certificate fingerprints, negotiated cyber suite).

2.4.6 Privacy preserving schemes

Privacy-preserving schemes can be used in two ways to detect energy theft attacks; one, focusing on protecting the identities of users, and the other, emphasising protecting the data of users (Guan et al., 2018). For the first aspect, pseudonym, anonymization, and virtual ring have been used. Pseudonym is considered to be a common user identity protection approach. The registration process for a pseudonym often involves many data protection methods, such as ring signature and zero-knowledge proof (Guan et al., 2018). Anonymizing smart-grid data is one of the methods approved by the National Institute of Standards and Technology (Afrin and Mishra, 2016). The main goal of anonymization is to enable smart grids' nodes to communicate in an anonymous manner with various smart-grid service providers by

using different pseudonyms. Another common method for user-identity preservation is a virtual ring, where a ring signature is used to validate the identity of users, without knowing their actual identity, by a control centre (Alladi et al., 2019). On the other hand, for the second aspect, emphasising protecting users' data, many methods can be used, such as data aggregation or authentication methods. Data aggregation is a well-known scheme which is used to protect the data of smart-grid users. It generally includes data obfuscation algorithms and homomorphic encryption (Guan et al., 2018; S. Salinas, Ming Li, and P. Li, 2012). Authentication methods are efficient countermeasures for privacy-related attacks and are usually based on key public infrastructure (Chin, Lin, and H.-H. Chen, 2016).

2.5 Summary

Smart power grids aim towards resilient, reliable and sustainable operation of legacy power systems and also the integration of smart business models for the optimised use of energy by consumers. Nonetheless, their complex system architecture in which diverse and heterogeneous infrastructures interconnect, facilitates the basis for a number of attacks that enable energy theft. Energy theft attacks affect critical grid processes and facilitate financial gain for malicious actors. To present the overall overview of such actors and their energy theft activities, we conduct a through study of different energy theft attacks and detection techniques in this chapter for smart grid systems.

In this regard, we firstly present the smart grid components in the energy supply chain with a focus on their data communication along with the pillars to access grid effectiveness. The impact of energy theft in the smart grid is then discussed by critically assessing how energy theft can be formulated by manipulating demand, supply, and generation data. The data-driven and data-agnostic energy theft attack examples are then discussed along with their enabling activities. Furthermore, we categorize extensive studies addressing the data-driven aspect of energy theft detection and summarizing the experimental approaches for such studies. Lastly, we highlight various open issues and challenges still persisting in the area of energy theft detection. We summarise and further indicate research directions for energy theft.

According to the study conducted in this chapter, there are still some uncertainties over the strategies and techniques used by malicious actors to perpetrate energy theft activities and manipulate the business model of modern energy sectors to gain financial benefits. Such uncertainties are extended to cover aspects related to the architecture, components, and resources of energy theft detection strategies. Therefore, it is crucial to utilise the current electricity market, which is driven by a demand for data collection and analysis, to design an overarching data-driven detection framework that enables

utility providers to develop energy theft detection methods based on a variety of theft attack scenarios. Motivated by this observation, the subsequent chapter tackles the aforementioned scenario and proposes a general theoretical framework for the data-driven process of detecting energy theft activities across smart grid infrastructures.

Chapter 3

Energy Theft Detection Framework

The smart grid is an automated, geographically diverse energy delivery network that integrates information and communication technologies to provide real-time information and enable instantaneous supply and demand balancing (Gellings et al., 2011). However, this system also exposes energy providers to new types of theft attacks, increasing financial losses (Qi et al., 2016; Sanjab et al., 2016). For instance, a Puerto Rico energy provider experienced systematic energy theft, costing around $400 million annually (Krebs, 2012).

Energy providers worldwide are reducing losses by detecting energy theft activities. Conventional detection methods, which involve physically inspecting locations where theft is expected to intensify, are time-consuming, inaccurate, costly and require significant human effort (Jindal, Dua, et al., 2016; Aldegheishem et al., 2021; Gao, Foggo, and Yu, 2019a). The smart grid paradigm offers data-driven detection approaches to conduct these inspections more efficiently (Gao, Foggo, and Yu, 2019b; K. Zheng et al., 2018). These data-driven techniques are distributed across various smart grid infrastructures, including generation, end-user infrastructures and transmission and distribution (T&D), and they can identify energy thefts perpetrated through various strategies, such as clandestine cyber connections or physical meter tampering (Jindal, Schaeffer-Filho, et al., 2020).

An analysis of studies on smart energy system security reveals that there is no standard data-driven method for detecting energy theft activities. In general, these methods employ various measurements gathered from the integrated data-collection infrastructure of modern energy grids. They adopt various approaches and algorithms from a broad spectrum of knowledge, with machine learning being the most prevalent (Messinis and Hatziargyriou, 2018). Therefore, in this chapter, we aim to take a significant step forward and abstract the energy theft detection process into a generically applicable theoretical framework from a data-driven perspective. The proposed theoretical framework provides a review of existing theories that serve as a

road map for developing a data-driven approach to the detection of energy theft in modern energy grids, commonly referred to as smart grids. As case studies for the application of the proposed theoretical framework to derive data-driven theft detection approaches, we also provide an overview of the data-driven detection approaches developed in this thesis. In summary, the contribution of this chapter is two-fold by providing:

1. An abstraction of the data-driven process of detecting energy theft activities in a general theoretical framework. From a data-driven perspective, this foundational model reflects commonalities in the phases, domains and dimensions that are involved in the energy theft detection process.

2. An overview of the two data-driven detection contributions made by this thesis, which are: i) a predictive energy theft detection approach for distributed renewable energy sources (DRES); and ii) an adaptive energy theft detection approach for consumption and DRES generation smart meters. Although the issue of energy theft in non-renewable energy sources is acknowledged, it lies outside the scope of this thesis.

The rest of this chapter is organized as follows: Section 3.2 discusses the proposed theoretical framework, while 3.3 discusses the operation dimension, which is the underpinning of our theoretical framework. Section 3.4 provides the overarching data flow of the energy theft detection approaches that contributed to this research, and Section 3.5 summarizes and concludes this chapter.

3.1 Energy theft detection requirements

Energy theft is a multidimensional problem, and multiple specific, tangible requirements should be considered when developing a data-driven detection strategy. These requirements identify the functionality required by a detection approach in order to maximise the return on investment in anti-energy theft initiatives. Fig. 3.1 illustrates these requirements. As depicted in this figure, the data-driven energy theft detection requirements are grouped into three main categories: must have, should have and could have. The requirements solicitation we conducted here was distilled using the MoSCoW method. It is adopted because it provides a consistent, high-confidence, low-complexity prioritisation process capable of managing multiple alternative inputs (Khan et al., 2015; Burgess and Sunmola, 2021).

The "must have" requirements specify the critical properties that a detection approach is required to possess, such as being independent of the functional components invoked. These properties are typically influenced by the nature of energy theft attacks and the constraints imposed by anti-theft initiatives. The "should have"

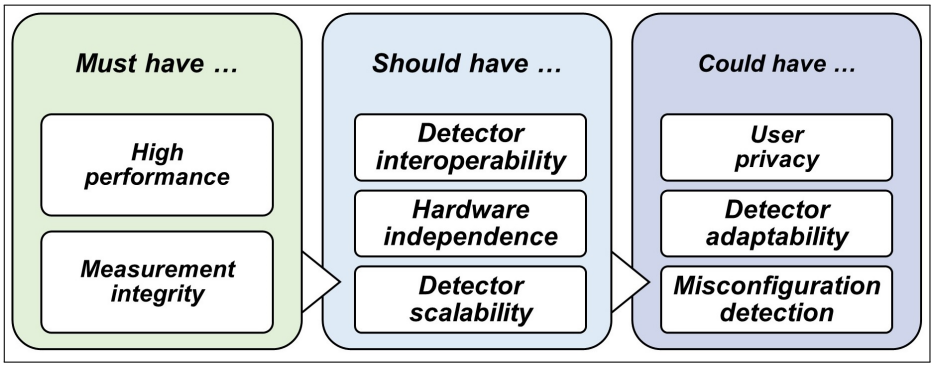

Figure 3.1: Requirements for data-driven energy theft detection.

requirements indicate the crucial, but not vital, properties of a detection strategy. Depending on how the detection method operates, these specifications might need to be satisfied or they might not. Hence, in the context of data-driven energy theft detection, the meeting of "should have" properties can be negotiated, whereas "must have" properties must be met. The properties labelled "could have" are not critical to the core functionality of the detection strategy. Hence, in comparison to the "should have" specifications, a failure to meet these requirements has less effect on the detection outcomes.

The requirements categorised as "must have" include high performance and measurement integrity. In this regard, the algorithmic properties of the detection solution should achieve a high level of accuracy and precision in identifying theft from streamed measurements from prosumers and consumers, with minimal computational requirements required to obtain classification decisions. In addition, because data-driven detectors rely on the energy measurements of grid users, which are vulnerable to adversarial attacks, the integrity of these measurements must be retained. Unfortunately, these adversarial attacks can compromise the robustness of energy theft detectors, such as a 17% drop in the detection rate (Takiddin, Ismail, and Serpedin, 2022; Takiddin, Ismail, Zafar, et al., 2020; Jingbo Hao and Tao, 2022).

The requirements categorised as "should have" involve detector interoperability, hardware independence and detector scalability. As the smart grid environment entails highly distributed domains, technologies, components, and infrastructures, the detection strategy should leverage a sufficient number of data sources to create a synergistic detection effect. Hence, obtaining data from multiple sources can generate a synergistic detection effect that maximises detection performance while being invulnerable to adversarial attacks (Krishna, Gunter, and Sanders, 2018; Glauner, Meira, Dolberg, et al., 2017; W. Li et al., 2019; Cody, Ford, and Siraj, 2015). In addition, the detection strategy should avoid dependence on the use of additional grid equipment. Many theft detectors (such as (Jindal, Dua, et al., 2016; Kim et al., 2019))

rely on additional hardware (such as observer metres) to detect theft. Nonetheless, deploying such additional resources imposes unnecessary costs. Finally, the detection strategy should be able to process large volumes of data in near real time in order to identify energy threats (S. Tan et al., 2017).

We also specify the "could have" requirements for the energy theft detection approach; these include user privacy, adaptability, and the detection of misconfiguration. Hence, the approach for detecting energy theft should maintain the confidentiality of user information, because the disclosure of private data could raise safety concerns and violate the users' privacy (e.g., if criminals gain access to this sensitive data, energy usage/generation patterns could be used to determine the daily routine of a household in order to commit crimes) (S. Salinas, Ming Li, and P. Li, 2012). In addition, the detection solution should be able to adapt and re-optimize its detection thresholds in response to the inclusion of new types of grid components and the installation of new technologies (e.g., the installation of energy-efficient (Fekri et al., 2021; Xia et al., 2022)). Typically, the deployment scenario for energy theft detection depends on the continuous arrival of energy readings. Hence, a theft detector could be incrementally adjusted by analysing energy instances as they become available in order to adapt to newly arriving energy patterns (Alkhresheh, Al-Tarawneh, and Alnawayseh, 2022). Finally, as a "could have" requirement, the algorithmic attributes of the detector could be developed so as to differentiate between theft-related behaviours and anomalous events that may have been caused by non-malicious intent (Yip, Wong, et al., 2017a; Jokar, Arianpoo, and Leung, 2016).

3.2 Theoretical framework of data-driven energy theft detection

In this section, we propose a framework that is tailored to serve the notion of energy theft detection and considers the requirements highlighted in Fig. 3.1. As shown in Fig. 3.2, the proposed theoretical framework for identifying energy theft activities in a smart energy system is a complementary four-dimensional structure. These four dimensions are infrastructure, measurement monitoring control (MMC), operation and end-user. These dimensions collectively constitute a dynamic structure that provides utility operators with continual updates on potential threats. By optimising data-driven detection procedures, utility providers can respond promptly, thereby minimising energy and monetary losses caused by energy theft threat. The framework essentially provides a proactive and resilient approach to securing intelligent energy systems to counter the constantly evolving attacks posed by energy theft.

The infrastructure dimension represents the endpoints utilised for energy theft activities. It comprises the physical deployment of energy grids, including generation

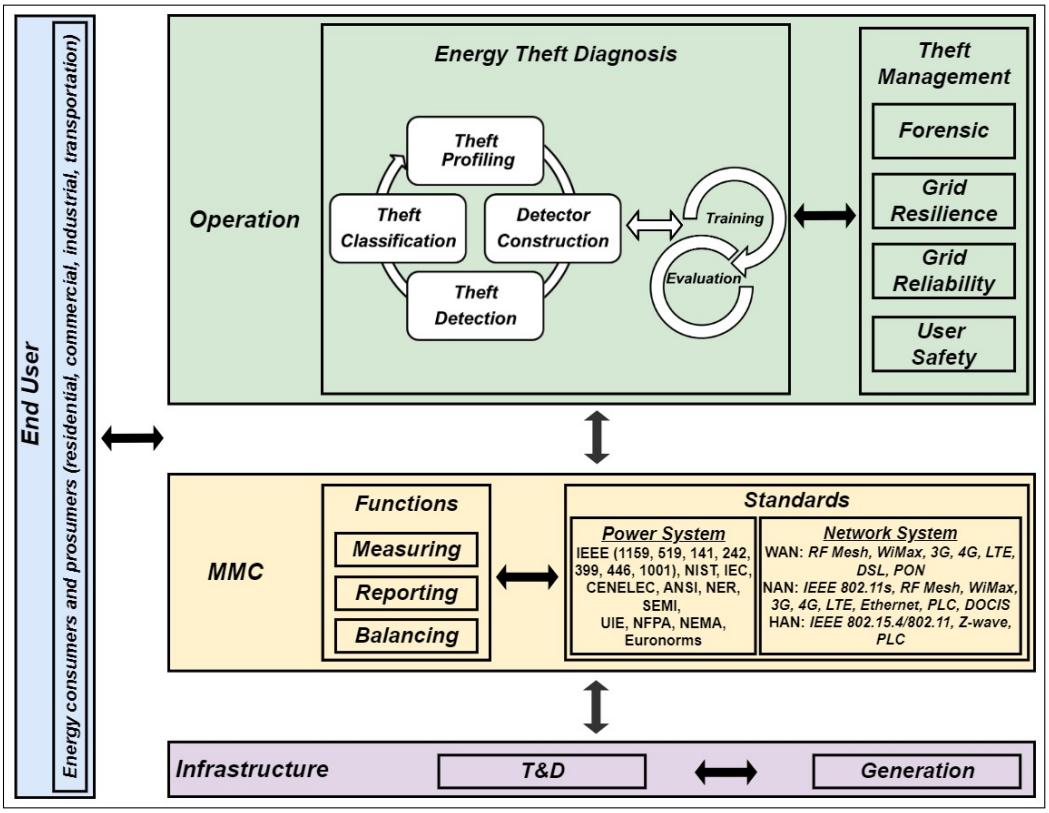

Figure 3.2: Theoretical framework for data-driven energy theft detection in smart energy systems.

facilities and T&D networks. The generation facilities involve centralised deployments (such as hydroelectric dams, wind farms, solar farms and fossil fuel power plants) and on-site renewables (i.e., DRES) such as wind and solar, and waste-to-energy. The T&D networks consists of all the electrical energy systems (low-voltage, high-voltage and medium-voltage), T&D stations, and transformers.

The MMC dimension of the proposed theoretical framework manages the data collection infrastructure integrated into modern energy grids. Therefore, it contributes in two distinct ways to the data-driven energy theft detection process. On the positive side, this process relies solely on the measurements collected by this dimension across energy systems. Nevertheless, this dimension creates more vulnerable endpoints that can be used to launch energy theft attacks. Essentially, it has three functions (measuring, reporting and balancing) that are governed by two standards (communication network standards and electric power system standards). These components rely on networked environments that inherently lack security procedures. Malicious actors exploit these vulnerabilities by manipulating

communication and energy measurement data produced and stored by networked measuring, reporting and balancing elements. By tampering with data integrity and energy measurement accuracy, these entities can present erroneous information, leading to financial gains through energy theft or fraudulent trading (Mahmoud et al., 2020; Jindal, Dua, et al., 2016; Messinis and Hatziargyriou, 2018).

However, the actual process of data-driven energy theft detection is conducted by utility providers in the operation dimension. This dimension obtains data and measurements from the MMC dimension that are used for detecting theft-related activities. As Fig. 3.2 illustrates, the actual process of theft detection is one component of a larger operation-level process: energy theft diagnosis. The output of this process is used by the theft management unit, which is responsible for utilising the detection process to reduce utility providers' financial and energy losses and to maintain the effectiveness pillars of the energy system.

The fourth and final dimension of the theoretical framework for energy theft is the end user dimension, which penetrates the previous three dimensions. As shown in Fig. 3.2, the infrastructure and MMC dimensions include residential, commercial, industrial and transportation energy consumers and prosumers. Their physical deployments are a crucial component of the energy system infrastructures, and their energy measurements represent the main part of the MMC. Furthermore, malicious grid users (prosumers and/or consumers) are the intrinsic actors involved in energy theft activities; their goal is to manipulate the business model and profit financially. Therefore, they represent an essential component of the operation dimension of the proposed theoretical framework for energy theft detection.

The operation dimension is discussed in greater depth in the following subsections as it is the central pillar of the proposed theoretical framework and contains the actual energy theft detection process.

3.3 Operation dimension

3.3.1 Energy theft diagnosis

The energy theft diagnosis process is essentially a cyclical process comprising four sub-processes: theft profiling, detector construction, theft detection and theft classification.

3.3.1.1 Theft profiling

During the theft profiling process, energy theft attacks are conceptualized. Essentially, the aim of the profiling process is to simulate energy theft events over smart grid components. It synthesises theft observations and feeds them into the detectors by

developing energy theft functions associated with fraudulent user behaviours and energy-measurement patterns. In this regard, the energy theft function accurately simulates various electrical configurations associated with energy theft. It does this by mimicking strategies and attack vectors that have a physical basis for their effects on energy-measurement patterns.

Primarily, there is a noteworthy absence of accessible (and applied) prototype implementation in practical theft situations, as well as a lack of data sets that allow thorough experimental analysis and are specifically suited to energy theft detection (Messinis and Hatziargyriou, 2018). Although there are a multitude of studies on energy theft attack detection (e.g. (W. Hu et al., 2020; Ramos et al., 2011; Meira et al., 2017; Jeyaraj et al., 2020; Massaferro, Martino, and Fernández, 2022)), their development is based on instances of historic theft. While they therefore do not need to fabricate theft observations, these studies are incapable of detecting new advanced and stealthy attacks that have no previous examples.

This is an issue because malicious actors continue to investigate new strategies and resources for perpetrating energy thefts across smart grid infrastructures (Krishna, Gunter, and Sanders, 2018). In turn, this has a negative effect on the overall detection performance of these approaches (Júnior et al., 2016; Aydin and Gungor, 2018). The majority of historical theft observations and their associated data sets are mostly prepared for specialised projects, limiting the generalisation of findings (Cintuglu et al., 2016). By profiling energy theft attacks, defenders obtain deeper insights into attackers' behaviours, enabling them to anticipate a spectrum of theft-related events and to develop effective and generic data-driven detection approaches.

The theft profiling process includes the following four procedures:

1. Endpoint assessments: this procedure identifies vulnerabilities in smart energy systems that can be exploited in theft attempts. The assessments entail scanning smart energy systems' cyber and physical resources to evaluate the likelihood of these resources being exploited by malicious actors for financial gain. Put simply, this procedure identifies vulnerabilities that enable adversaries to exploit grid infrastructure components, communication networks and applications, with the intention of stealing money.

2. Strategy identification: this procedure tackles the question of how vulnerabilities are utilized to perpetrate thefts. Understanding this entails identifying the methods, techniques and attack vectors that thieves use against each vulnerability to achieve their goals.

3. Data acquisition: the purpose of this procedure is to use the MMC dimension to gather measurements from the identified vulnerabilities. As Fig. 3.2 illustrates, the MMC dimension employs different communication standards and protocols

for data acquisition. For instance, proprietary protocols (Z-wave, ZigBee) gather measurements from home area networks, while WiMax and IEEE 802 series gather data from T&D sectors (Ma et al., 2013).

4. Theft fabrication: this procedure uses the findings from the endpoint assessments and strategy identification to generate profitable energy theft attack functions. These can be applied to the findings from the data acquisition procedure to produce realistic synthetic fraudulent behaviour patterns. These are injected into legitimate measurements to explore and greatly enhance the energy theft detection process.

3.3.1.2 Detector construction

The detector construction process adopts a methodology capable of learning the complex, non-linear patterns within the data prepared in the theft profiling process. This acquired knowledge can be subsequently employed to distinguish between legitimate observations and theft attempts. Primarily, this process generates a detector, which comprises a variety of algorithms and methods that are trained to detect theft by identifying patterns in the data from the profiling process (i.e. the synthetic theft observations and acquired legitimate measurements). The detector implements a set of instructions to provide data-driven predictions and decisions regarding the presence or absence of energy theft attacks in smart grid infrastructures.

The diverse algorithmic proprieties within the detector construction process can address the challenges of learning energy patterns and identifying possible fraudulent actors over smart grid infrastructures. For instance, data-level machine learning techniques (e.g. SMOTE and condensed nearest neighbour rule) are used to address the challenges of an imbalanced data category, which occurs when one category (either fraudulent or legit instances) contains significantly more samples than others (Krawczyk, 2016). Moreover, the differences, averages or other statistical transformations of the original feature vector can be used to construct and refine additional features that enhance the effectiveness of the theft detection methodologies (Glauner, Meira, Valtchev, et al., 2016; Maamar and Benahmed, 2018; Jundong Li et al., 2018). In addition, a meta-learning scheme (i.e., a learning process involving the collection of information from previous experiences, to be utilised for future learning) might reduce the false positive rates caused by non-malicious activity (Jinghang Li and M. Hu, 2020). Finally, a binary classifier-based intrusion detection system trained on energy device behaviour logs can be employed to combat adversarial machine learning (Bor et al., 2019). On the other hand, to preserve the security and privacy of the collected energy measurements, a federated learning-based technique can be employed (M. Wen et al., 2021; Ashraf et al., 2022).

As depicted in Fig. 3.2, the detector construction process entails two intrinsic

procedures: training and evaluation. During the training procedure, the selected model learns patterns within the data that were prepared during the theft profiling procedure, in order to acquire knowledge that can be used to detect energy theft. The training procedure entails four processes: data cleaning, feature selection, dimension reduction, and model selection and tuning.

During the data cleaning process, duplicate, erroneous or otherwise improper data are eliminated prior to the training procedure. Despite their similarities, feature selection and dimension reduction affect the input variables of the training procedure in a completely distinct manner. Feature selection is the process of selecting what characteristics to include or omit from consideration without modifying them, while dimension reduction transforms the input variables into a lower dimension. The model selection process involves choosing the most appropriate model for the energy theft detection process from a set of machine learning candidates. Finally, model tuning attempts to find the optimal values of the hyperparameters of the selected machine learning model, as the detection performance is maximized.

During the evaluation procedure, the learned model predicts categories (labels) of instances that were not introduced in the training procedure. The model outputs tagged data that predict whether the instances are malicious or legitimate. Consequently, in this research, the prediction outputs were compared with the correct categories of these examples, thereby quantifying the detection approach's overall performance.

In this regard, the evaluation performance metrics of the classifications are applied; these include the confusion matrix, accuracy (ACC), area under the curve (AUC), precision (PR), recall (RE), sensitivity and F1-score (F1). When the real category (malicious or legitimate) is known, the performance of the detection model can be described using a table called the confusion matrix. ACC measures how reliably the trained model can distinguish between malicious and legitimate instances. The model's ability to determine whether an instance is malicious is quantified by its AUC. PR is the proportion of malicious incidents correctly identified as malicious from all maliciously predicted observations, whereas RE is the proportion of malicious activities correctly identified as malicious from actual malicious samples. F1 provides a holistic perspective of PR and RE.

Generally, understanding the energy dynamics of data-driven algorithm and machine learning training within the detector constriction process involves examining multiple influencing factors, including the size of the data set and the complexity of the detection model. Without specific implementation details, the precise quantification of energy consumption remains challenging, but existing studies (such as (Yang Liu and S. Hu, 2015; Gunturi and Sarkar, 2021)) highlight the promise of optimised machine learning models with relatively low energy consumption for energy theft detection.

Despite the energy consumed by machine learning in the detector construction process, energy theft causes substantial MWh losses. For a single energy provider in Canada, such thefts result in an average annual loss of 850000 MWh, which converts to a monetary loss of $55 million (Raggi et al., 2020). Moreover, a Brazilian electricity regulatory agency estimates that 5% of the energy injected into distribution grids is lost due to theft activities (Carr and Thomson, 2022). A recent study reveals that nearly 20% of India's total generation of electricity is lost due to theft instances (Razavi and Fleury, 2019).

While acknowledging the energy consumption associated with machine learning, the potential benefits, such as reducing the need for physical inspections to detect potential energy theft in locations where energy theft is expected to increase, suggest that employing data-driven detection techniques facilitated by machine learning techniques may be a worthwhile investment to reduce the significant MWh losses caused by energy theft attacks.

3.3.1.3 Theft detection

After completion of the detector contracting procedure, the constructed detector is ready for theft classification. The detector will be used to detect real-world energy theft attacks across smart grid systems. This makes the detector's predictions of the legitimacy of prosumer and/or consumer energy measurements available to utility providers, allowing them to make data-driven decisions over energy theft activities. Consequently, the theft detection process is the culmination of the measurements collected from the infrastructure and end user dimensions through the MMC dimension, and profiled in the operation dimension for use in the detector also constructed in the operation dimension.

3.3.1.4 Theft classification

It is important to note that the outcome of the theft detection process correlates with predictions of anomalies that may be caused by non-malicious activities (e.g., smart-meter misconfigurations). In fact, this is a classic issue within the data-driven energy theft detection process. These non-malicious anomalies reduce the efficiency of the theft detection process. When these variables are not effectively addressed, they can cause a significant number of false positives, in which consumers and/or prosumers who were identified to be malicious were actually legitimate (Jokar, Arianpoo, and Leung, 2016). Such detection outcomes are expensive for utility providers, as they need costly, time-consuming, and labour-intensive on-site inspections.

The theft classification process aims to resolve this issue by classifying the anomalous events into various categories. During this procedure, the distinctions between malicious and legitimate processes are identified, and the specific properties

of energy theft incidents and misconfiguration events are highlighted for use in the subsequent theft profiling process. This then generates adequate synthetic theft and misconfiguration observations.

This step entails establishing an accurate profile of energy theft and misconfigurations within the energy system. It entails all processes of theft profiling in Section 3.3.1.1, including endpoint assessments, strategy identification, data acquisition and theft/ misconfiguration fabrication, with a focus on identifying patterns or anomalies associated with theft attacks or system misconfigurations.

Once we have a comprehensive understanding of these patterns, we can use them to generate synthetic theft and misconfiguration observations. In other words, based on the characteristics identified, we can generate simulated instances that mimic real-world incidents. These synthetic observations are indispensable for testing and refining detection algorithms and security measures, thereby ensuring that the effectiveness pillars of the energy system are maintained.

3.3.2 Theft management

Within the proposed theoretical framework for energy theft, the theft management unit leverages the output of the energy diagnosis process to mitigate the cascade impacts resulting from theft attacks on energy systems. The attacks underpinning energy theft have a negative impact on system optimization procedures. In this regard, the energy diagnosis procedure is performed cyclically to support the energy management unit in mitigating the negative impact of a theft attack on the smart grid efficiency and business model.

By virtue of the theft management unit within the operational dimension, on-site inspections are only required for final verification after an energy theft attack has been identified, and they thus improve the overall grid forensics (Jokar, Arianpoo, and Leung, 2016; K. Zheng et al., 2018). Furthermore, the diagnosis of energy theft maintains grid reliability. When energy theft occurs, the power supply system is overloaded and unable to meet customer demands (Ahmed et al., 2022). Consequently, this overloading is mitigated and the demand-supply imbalance is managed by detecting energy theft activities.

Moreover, energy-theft-related attacks against T&D systems can cause state estimation deviation. This misguides decisions, potentially negatively affecting the resilience of the smart grid (Xue, Jing, and H. Liu, 2019). Such detrimental impacts can be avoided by detecting theft attacks. Finally, grid user safety can be preserved as a result of the energy diagnosis output because theft-related manipulation of grid components, which that can cause electrocution or fires, leading to property destruction and even death, can be detected and mitigated (Czechowski and Kosek, 2016).

3.4 Energy theft detection data flow

This thesis proposes two principal data-driven energy theft approaches as case studies derived from the theoretical data-driven energy theft detection presented in Section 3.2. Note that the implementation of both proposed strategies for energy theft is entirely dependent on Python. We utilised established built-in libraries such as TensorFlow and Scikit-Learn, but made all necessary modifications to ensure they meet the requirements of our proposed energy theft detection methods. These modifications required extensive adjustments and adaptations to the functions and classes within these libraries, allowing for their complete incorporation into our codebase. An overview of data flow in these detection approaches is presented in Fig.3.3.

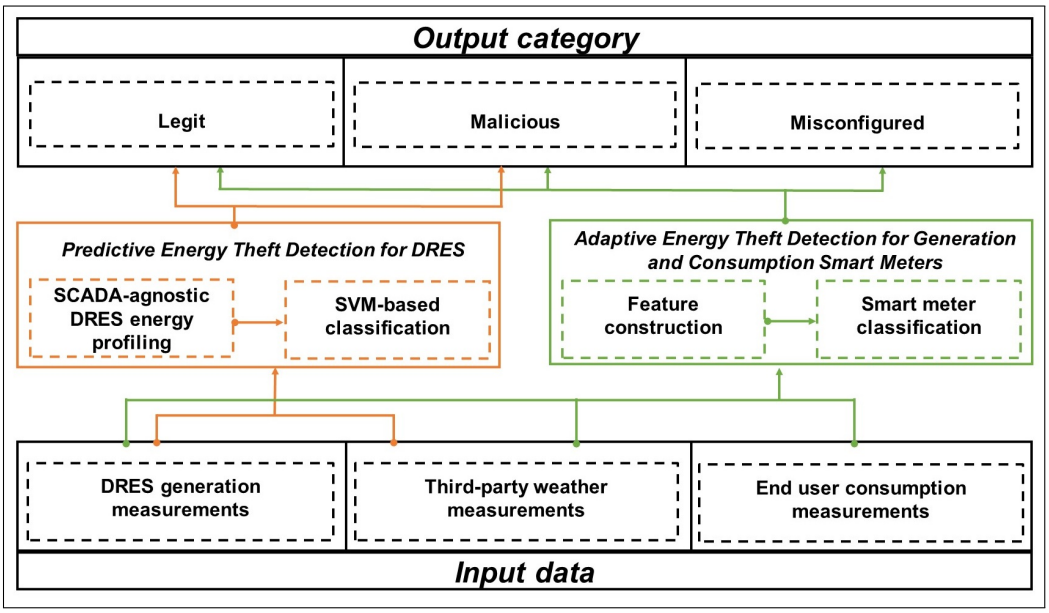

Figure 3.3: Overview structure of the proposed data-driven energy theft detection approaches.

As depicted in Fig. 3.3, the first approach, predictive energy theft detection, explicitly detects energy theft in DRES-based infrastructure. Keep in mind that energy theft detection in non-renewable energy sources is beyond the scope of interest of this thesis. The second approach, adaptive energy theft detection in smart grids, is a generic approach that detects theft in DRES generation and end-user consumption infrastructures. The development of these detection methods begins with the operation dimension of the proposed theoretical framework, specifically within the theft profiling sub-process of the energy theft diagnosis process, depicted in Fig. 3.2.

According to the theft profiling sub-process of the proposed theoretical framework, vulnerabilities in smart energy systems that could be exploited in energy theft attempts should be identified initially. Hence, in the development of our first proposed detection method, we demonstrated that the DRES installations owned by prosumers who are not operators of power transmission or distribution networks represent vulnerabilities that are exploited for financial gain by malicious actors. DRES operating costs are approximately US$15/MWh, which exceeds the range of US$0 to US$45/MWh paid to DERS owners (Krishna, Gunter, and Sanders, 2018). This creates a motivation for malicious prosumers to maximise their reported generation fraudulently in order to maximise their profits (Mahmoud et al., 2020).

Such energy theft can be perpetrated using a variety of methods identified during the strategy identification procedure within the theft profiling sub-process of the proposed theoretical framework. To maximise the reading of DRES smart metres, the identified DRES-related theft method includes a physical scenario (i.e., the use of a solar array simulator) or a cyber scenario (e.g., accessing the firmware of the measuring systems) (Shaaban et al., 2021; Yuan, M.-g. Shi, and Sun, 2015).

The MMC dimension of the proposed theoretical framework is then employed to collect generation measurements from legitimate DRES deployments for use in the theft fabrication procedure within the theft profiling sub-process. Four functions of theft scenarios were identified during the development of our first detection approach. These four functions mimic actual fraudulent patterns in terms of reporting erroneously generated energy. Then, by developing the proposed energy theft functions, synthetic anomalies are injected into the legitimate datasets to create a dataset containing both legitimate and fraudulent patterns.

The resulted dataset is then used in the detector construction sub-process of the operation dimension of the proposed theoretical framework. As depicted in Fig. 3.3, two data-driven algorithms are adopted here to accomplish the theft detection process: i) a supervisory control and data acquisition (SCADA)-agnostic DRES profiling approach working solely on third-party and widely available weather observations; and ii) a classification scheme relying on DRES profiling that is capable of classifying theft detection activities. After completion of the contracting process, this detector is ready for theft detection. The output of our first theft detector is a label for each smart meter of each DRES, indicating whether it belongs to a legitimate or malicious prosumer. In Chapters 4 and 5, we describe the development of our first proposed detection method, titled adaptive energy theft detection in smart grids, in detail.

As the energy theft diagnosis process within the operation dimension is a cyclical process that can be continually optimised, we use the feedback from our first approach, predictive energy theft detection, to develop our second approach, adaptive energy theft detection for generation and consumption smart meters. Although our first detection strategy yielded a satisfactory performance, it was limited in its scope as

it considered energy theft in DRES supply measurements only. Hence, it is unable simultaneously detect theft activities in generation and consumption measurements.

This strategy imposes a significant computational expense on the utility provider, as it requires them to simultaneously apply two distinct approaches for detecting theft in both measurements. Consequently, it necessitates additional computational resources for training, evaluating, and deploying these distinct approaches. In parallel, although our first detection method focuses on a single measurement, it is incapable of distinguishing energy theft from anomalous events caused by legitimate incidents (i.e., smart metre misconfiguration). Finally, our proposed first detection approach fails to adapt and re-optimize its detection thresholds as it should, in light of the addition of further grid components and the implementation of new technologies. In fact, the majority of studies on data-driven energy theft contain all of these gaps and challenges.

Therefore, during the development of our second detection approach, we extend the theft profiling sub-process of the proposed theoretical framework and propose a general adversary model applicable to stealthy energy theft and abnormalities caused by legitimate events in consumption and generation measures. Hence, during the procedure of strategy identification, we identify a number of methods that malicious actors can utilise to maximise the readings of their DRES generation and/or to reduce the readings of their consumption. In parallel, we differentiate the patterns of these malicious methods on the consumption and generation measurements from anomalies that result from grid component misconfigurations. As a result of these procedures, we introduces a taxonomy of smart metre anomaly functions based on energy consumption and generation measurements that can be used to fabricate theft and misconfiguration samples.

Consequently, we employed the output of this round of the theft profiling and continue the detector construction sub-process within the operation dimension of the proposed theoretical framework. In this regard, we propose a combination of an adaptive feature constriction method and a smart energy meter classification component for constricting theft attack detectors. The proposed approach evolves from an aggregation of weather observations, and theft and misconfiguration events across DRES and consumption installations. In addition, it can self-optimize based on incoming measurement stream properties without human intervention. The output of our second theft detector is a label for each smart meter of each prosumer and/or consumer that indicates whether the smart meter is legitimate, malicious, or misconfigured. In Chapter 6, we describe in detail the implementation of our second proposed detection method, titled adaptive energy theft detection for generation and consumption smart meters.

3.5 Summary

This chapter presents a generally applicable theoretical framework of a data-driven process for the detection of energy theft. The framework is developed based on understanding the energy theft problem in the depth through an appropriate requirement for use with modern energy grids. Since there is currently no standard data-driven approach for identifying energy theft activities, this chapter aims to make a significant advance and abstract the process of detecting energy theft into a theoretical framework that is generically applicable from a data-driven perspective. We also provide an overview for the data-driven detection approaches developed, implemented and evaluated in the following three chapters (i.e., Chapter 4, Chapter 5 and Chapter 6) as case studies of the proposed theoretical framework's application, with the aim of developing data-driven theft detection techniques.

Chapter 4

SCADA-agnostic Energy Modelling for Distributed Renewable Energy Sources

The use of fossil fuels for power generation has led to alarming air pollution and carbon emission levels that consequently impact negatively to climate change and global warming. According to the international energy agency (IEA), there was a 2.3% rise in energy consumption just in 2018, which caused CO_2 emissions to rise by 1.7% leading to an alarming value of 33.1 Gt of CO_2 in the air (IEA, 2018). Hence, modern smart grid deployments adopt greener power generation solutions based on distributed renewable energy sources (DRES) including Photo Voltaic (PV) solar panels, wind turbines and bio fuel. DRES deployments are expected to take a significant portion of the global energy generation reaching 40% of the smartgrid ecosystem by the year 2020 (IEA, 2020).

Nonetheless, as the DRES generated power output depends solely on intermittent environmental conditions (e.g., ample solar radiation, wind speed), there is always a level of uncertainty in terms of the power contribution that such deployments offer back to the main power grid. Under the objectives of a sustainable grid, it is therefore crucial to adequately profile and further forecast DRES power production. Grid optimisation routines rely on accurate DRES profiling, thus inaccurate and unavailable DRES profiling is highly likely to trigger resilience havoc with a number of severe consequences.

Power generation profiling for DRES has been the subject of investigation in a number of studies (e.g., (Janssens et al., 2016; Pelletier, Masson, and Tahan, 2016)). The majority of these studies engage with the assumption that measurements from Supervisory Control and Data Acquisition (SCADA) systems are always available and the various modelling components are restricted on explicitly utilising power

generation values from such systems. For instance, (Y. Wang et al., 2018) relies on wind speed timeseries and employ a spline regression model to model power generation of wind turbine deployments for power forecasting. The work in (Abedinia et al., 2020) suggested a hybrid framework dedicated at wind power profiling based on empirical mode decomposition in conjunction with a bagging neural network, and a stochastic optimization algorithm. However, the complex business and operational processes within large-scale power grids involve a diversity of ownership in terms of machinery (e.g., wind turbines) as well as control and measurement components (e.g., PLCs, SCADA) (Leahy et al., 2019; Ahmed et al., 2022; X. Cheng et al., 2022). Thus, the acquisition of SCADA-based measurements is not always available, particularly for DRES installation owners that are not operators of either power transmission or distribution networks.

In this piece of work, we tackle the aforementioned scenario and propose a generic SCADA-agnostic DRES power profiling scheme. To the best of our knowledge, no other studies have considered the pragmatic assumption that SCADA measurements and sensor components are not present or available in all DRES deployments. The proposed DRES profiling system is an instantiation of the data-driven theoretical framework presented in Chapter 3, paving the way for the independent detection of fraudulent activities underlying energy theft. It makes use of freely available weather measurements through the MMC dimension of the proposed theoretical framework in order to explicitly profile and predict the power generation produced by actual wind turbine deployments across the infrastructure dimension. Then, in the operation dimension of the proposed theoretical framework, the proposed generation profiling approach enables the automated feature selection and tuning of machine-learning-based regression models. These models are capable of operating adequately with freely available third-party weather measurements for developing a SCADA-agnostic detection process explicitly for DRES-based theft scenarios.

In general, the main contributions of this chapter are two-fold and summarised as follows:

- A generic DRES profiling system enabling adaptive feature selection as well as automated best-fit machine learning model tuning under low computational costs.
- A SCADA-agnostic cost-efficient approach relying strictly on freely available third-party weather measurements to model DRES deployments with a proof-of-concept evidence over real wind turbine deployment profiling.

The rest of this chapter is organised as follows. Section 4.1 describes the datasets and the methodology of the proposed approach whereas Section 4.2 discusses the evaluation conducted. Finally, Section 4.3 concludes and summarises this chapter.

4.1 Data Description and methodology

4.1.1 Data description

The herein reported proof-of-concept study focuses explicitly on profiling wind turbine deployments using third-party, freely available weather measurements under the assumption that SCADA or locally placed sensor-based measurement data is not available. However, in order to validate the performance of our exogenous-based wind power modelling, we utilise SCADA measurements gathered from a real deployment.

The used SCADA-based dataset was captured at the La Haute Borne wind farm, located in Meuse, France[1] and represents daily measurements gathered for the whole year of 2017[2]. The La Haute Borne wind farm consists of 4 Senvion MM82 wind turbines where measurements are obtained on 10-minute samples. Within each 10-minute sampling bin, there are 34 features related to various electro-mechanical (e.g., torque, rotor speed), power (e.g., apparent power, grid voltage) and environmental parameters (e.g., wind speed, outdoor temperature) explicit to a given wind turbine.

As already mentioned, our SCADA-agnostic scheme depends solely on third-party weather measurements that are freely available. For this purpose, we have extracted environmental measurements (e.g., wind direction) from the Dark Sky API[3] and Weather Online API[4] over the same observational period in which ground truth SCADA measurements were obtained for the La Haute installation. Moreover, we acquired wind and output temperature measurements from Weathernews[5] as observed by the Nancy-Ochey weather station which is geographically adjacent to the La Haute Borne wind farm. In total, our SCADA-agnostic dataset has 42 weather features including numerical and categorical data with in 1-hour sampling bin.

In order to achieve data synchronisation and maintain consistency, we organised SCADA-based measurements and third-party weather measurements on an hourly basis. The synchronised datasets were then merged into two consolidated data frames, one for the SCADA-based measurements and the other for the environmental measurements collected by third parties. Each data point in these data frames is identified by a distinct timestamp. Both the third-party weather measurements as well as the SCADA-based measurements were processed within our generic DRES profiling system that we explain next.

[1]Explore – ENGIE France Renewable Energy Open Data, Available: `https://opendata-renewables.engie.com/pages/home/`

[2]The 2017 dataset is the most complete in comparison with all datasets for other years provided by ENGIE.

[3]Explore – Dark Sky API, Available: `https://darksky.net/dev`

[4]Explore – World Weather Online API, Available: `https://www.worldweatheronline.com/developer/api/`

[5]Explore – Weathernews, Available: `www.weathernews.fr`

4.1.2 DRES profiling system

This study relies on a system built to efficiently pre- and post-process DRES measurements such as to automatically identify the most suitable features within a best-fit machine learning model. The generic properties contained within our implemented system can serve the basis for close-to-real-time profiling of any type of DRES deployment (e.g., wind turbine/farm, solar PV panels etc.).

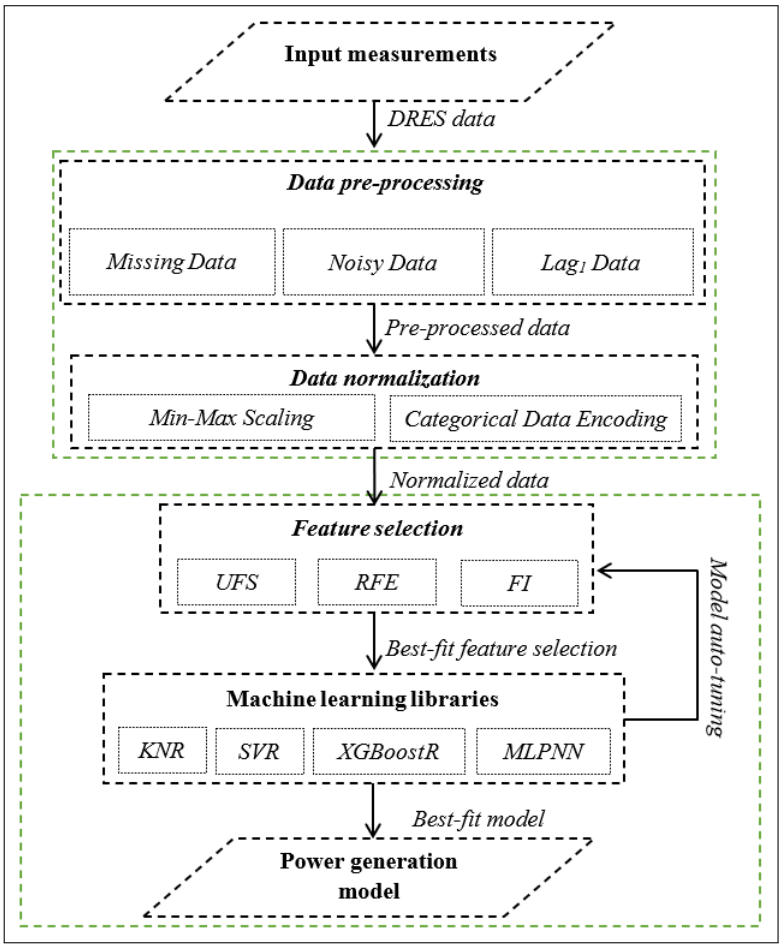

Figure 4.1: Measurement-based DRES profiling system.

As depicted in Fig 4.1, the first process within the implemented system is to pre-process diverse DRES measurements gathered either by conventional SCADA or sensor-based data acquisition deployments. Hence, the pre-processing module ensures that raw timeseries measurements of various features (e.g., wind speed, humidity, output power, etc.) are refined in terms of missing values, noisy timeseries and (re)sampling. Subsequently, the system normalises the pre-processed timeseries

and feeds them directly to a feature selection software component that works in synergy with a machine learning component. Ultimately, the combination of the best statistical features alongside the best-fit model is chosen based on a repetitive auto-tuning process. We describe the mechanics of each individual stage and component by focusing on the proof-of-concept wind power modelling scenario as follows.

4.1.2.1 Data pre-processing

The missing, duplicated and inconsistent samples caused by turbine unavailability, electrical shut-down, icing events, etc. can affect the accuracy of power measurement estimation. Therefore, the third party weather measurements and the power generation measurements are subjected to a filtering so as to remove all possible duplicated, missing and inconsistent (out-of-range) data samples. During the pre-processing stage, analysis on power measurements is performed through the autocorrelation function (ACF) and partial autocorrelation function (PACF) in order to build an underlying statistical ground-truth of the assessed timeseries (i.e., to extract the optimal lags (historical) features, denoted by Lag_n). Essentially, lags express the similarity of frequency components whilst treating power measurements as a signal in the function of time. Hence, the ACF represents the correlation between the P_t^{avg} (average generated power) measurement in $t \in T$ and the measurements at previous time lags. The PACF is the correlation between P_t^{avg} and P_{t+k}^{avg} after removing the influence of the confounding variable:

$$P_{t-1}^{avg}, P_{t-2}^{avg}, ..., P_{t-k+1}^{avg} \tag{4.1}$$

4.1.2.2 Data normalization

Our DRES profiling system employs a min-max normalisation scheme such as to reconstruct the assessed timeseries in the range $[0, 1]$ with $n \times m$ vectors as given in Eq. (4.2). In this case, n is the number of the samples, m is the number of feature vectors and $t \in T$ is a time interval.

$$\bar{x_t} = \frac{x_t - x^{min}}{x^{max} - x^{min}} \tag{4.2}$$

where $\bar{x}_t$ represents the normalized value of x_t, x^{min} and x^{max} are the minimum and maximum value in each feature vector $z \in m$ respectively. The normalization procedure is only applicable to the numerical features. As the dataset consists of a mixture of numerical and categorical features, the latter are encoded using the binary encoder function as follows: An integer value is assigned to every unique category for a given categorical feature. A new binary feature is created for each integer-encoded category and new columns are created based on the majority of the bit encoding. As

binary encoded features only take binary values of 0 or 1, they are not needed to be re-scaled or normalized.

4.1.2.3 Feature selection

The proposed DRES profiling system employs an automated feature selection process such as to obtain an adequate and effective set of attributes. Hence, the feature selection component is in charge of assessing the importance of the raw SCADA or third-party weather measurements. As evidenced in Fig. 4.1, the feature selection component works in synergy with the machine learning component such as to identify the optimal set of features producing the best-fit machine learning-based power regression model. In more detail, the current prototype supports: i) filter-based univariate feature selection (UFS), ii) wrapper-based recursive feature elimination (RFE) and iii) ranking-based feature importance (FI).

The UFS technique is used to assign the importance scoring of each feature. Thus, each feature is linearly regressed and produces an estimated value that is scored against the original value under the F-score metric. Essentially, the F-score denotes how the regressed value of a given input behaves in terms of the averaged accuracy precision. Our current prototype supports both the univariate linear regression filtering of features as well as filtering through the ranking of correlations based on the Pearson correlation metric. Both filtering mechanisms are used interchangeably. By contrast with UFS, the RFE method recursively selects features by removing the less important features from the feature set using importance-based rankings. Our current prototype utilizes the random forest (RF) estimator for importance-based rankings. Within the FI approach, a similar RF-based feature reduction is performed such as to isolate the most significant attributes. It is to be noted that both RFE and FI use RF to remove the least significant features; however the FI in contrast with the RFE is less robust as it is just based on a given threshold value and a single iteration.

4.1.2.4 Machine learning component

The implemented DRES profiling system depends heavily on the collaborative functioning between the feature selection component and the machine learning component. The machine learning component is implemented under a pluggable fashion in which off-the-shelf or customised machine learning algorithms can inter-operate with the algorithms residing within the feature selection process. The synergy between the aforementioned components is orchestrated under a repetitive feedback mechanism such as to identify the most optimal combination of features with an identified machine learning-based profiling model.

Moreover, optimal hyper-parameters for the machine learning-based techniques employed are found by using a grid search technique with a k-fold cross-validation

method. In detail:

1. Specify the hyper-parameters along with the value range for each model.
2. For each combination of hyper-parameter values in the specified grid:
   - Train the model with k-fold cross-validation. In this chapter, k is set to 10, and the model is trained and validated k times. Each time, a different fold is used as the validation set, and the remaining folds are used for training.
   - Calculate the average performance for each fold.
3. Identify the set of hyper-parameters that achieved the highest average performance in the grid search.
4. Train the final model using the entire training dataset and the optimal hyper-parameter values.
5. Evaluate the final model's performance on the test set.

This approach provides a systematic and efficient method for investigating a number of hyper-parameter combinations, ensuring that the performance of the model is systematically evaluated across multiple subsets of the data sources (Jindal, Kumar, and M. Singh, 2020b).

In order to address aspects of non-linearity in the examined features as well as properties of non-stationary DRES measurements, we have implemented both supervised as well as unsupervised machine learning-based regression algorithms. In particular, the current prototype supports: i) K-nearest keighbours regression (KNR), ii) support vector regressor (SVR), iii) gradient boosting regressor (XGBoostR) and, iv) multi-layer perceptron neural network (MLPNN). We next describe the basic properties of each implemented algorithm.

KNR: The KNR model utilises feature vector similarity (or neighborhood) and predicts the value of new input samples. Thus, the value assigned to new input samples is based on the resemblance with training samples. In summary, KNR is decomposed into three main stages;

1. Calculation of the Euclidean distance between the new input data instance with each training samples given by:

$$D_t = \sqrt{\sum \mid x_t^{train} - x_t^{new} \mid^2} \tag{4.3}$$

   where x_t^{train} and x_t^{new} represent the values of training sample and the new input data respectively.

2. k nearest samples are selected based on the closest Euclidean distance values.

3. Inserting the average of the k-nearest points as the predicted value of the new input instance.

SVR: The SVR model is a supervised scheme enabling the estimation of a fit function based on pre-computed training samples such as to map high-dimensional model inputs to the target output. Unlike other regression algorithms focusing on prediction error rate reduction, SVR fits any prediction errors within a a tolerable error (ϵ). Hence, describing the highest deviation from the targets, while keeping the fit function as flat as possible.

XGBoostR: The XGBoostR algorithm relies on the boosting idea is aiming to improve the regression stability of a weak learner that promote weak statistical hypotheses related to their input data instances. In general, a weak learner represents models holding slightly better performance than a random chance with respect to prediction error rates. XGBoostR depends on three components performing: i) loss function optimisation with respect to regression errors, ii) weak learner prediction for one decision at a time and, iii) weak learner additive model minimising the total loss function.

MLPNN: The MLPNN algorithm belongs in the category of supervised feed-forward artificial neural network (ANN) formulations and consists of more than one perceptrons. The input layer in MLPNN is used to receive input data, whereas the output layer is responsible for predicting the output value of a given input. Internally, the composition of the training model within MLPNN is performed by a back-propagation scheme. As within a traditional artificial neural network, hidden layers reside between input and output layers which work as computational engines.

In particular, MPLNN exploits the correlation or dependencies between the variables used in the computed training to model the output value by tuning weight parameters such as to reduce prediction errors. The number of the hidden layers is determined by applying the Hecht-Nelson method meaning that the size of the hidden layer $= 2n + 1$, where n is the size of the input layer (C. Ren et al., 2014).

4.1.3 Evaluation methodology

We conduct a thorough evaluation in order to assess the performance of the exogenous SCADA-agnostic wind power modelling in comparison with modelling performed using SCADA-based measurements. Our evaluation methodology is diagrammatically depicted in Fig. 4.2.

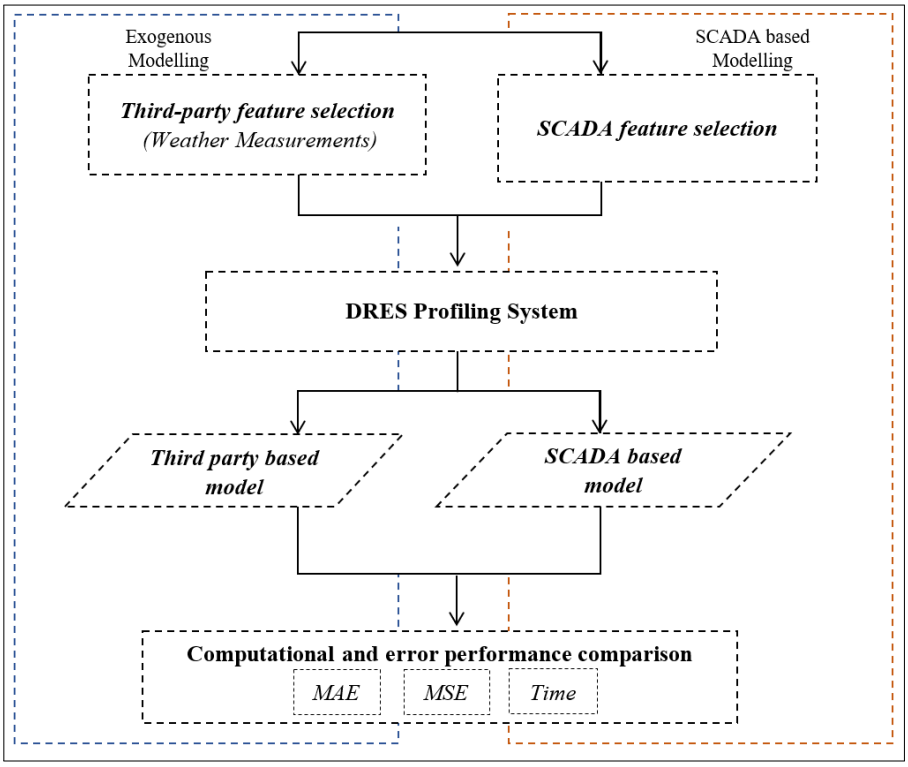

Figure 4.2: Evaluation methodology.

Both SCADA and SCADA-agnostic data streams are passed through our DRES profiling system prototype to obtain the most optimal features with the best-fit regression models. Prior to the modelling as well as the feature selection phase, the correlation of individual generated power with its past measurements is extracted and integrated to the input measurements of the designed system. Subsequently, we perform a seasonality grouping for every type of measurements for better classification. Hence, we split our datasets in the four seasons of the year (i.e. spring, summer, autumn, and winter) for each wind turbine and re-sample the measurements to behave under hourly bins. The DRES profiling system assigns 70% of the feature samples to be used for training for any of the algorithms within the machine learning component and 30% for testing.

Subsequently, the repetitive process between the feature selection component and the machine learning-library component takes places such as to identify the most optimal features for the best-fit model. The resulted models are assessed based on two error and one computational cost metric. The indices considered in this work in terms of prediction error are the mean absolute error (MAE) and the mean squared error (MSE), whereas for computation, we account the time taken to obtain a prediction.

We briefly describe each metric as follows.

1. **MAE:** The mean of all absolute values of the difference between the actual and predicted power values defined as:

$$MAE = m^{-1} \sum_{t=1}^{m} \mid x_t - \hat{x}_t \mid \tag{4.4}$$

   where $t \in T$, m is the test set length, $x_t, \hat{x}_t$ represent the actual power measurements and the estimated power measurements, respectively.

2. **MSE:** The mean of the squares of all differences between the actual and predicted powers defined as:

$$MSE = m^{-1} \sum_{t=1}^{m} (x_t - \hat{x}_t)^2 \tag{4.5}$$

3. **Computational complexity:** Time taken by the machine learning-based model within the DRES profiling system to produce prediction for the output power of a given wind turbine.

The implementation of the machine learning component and feature selection algorithms in our entire integrated system relies entirely on Python. We utilised standard built-in libraries such as TensorFlow and Scikit-Learn, but with significant customization and adaptation to align them with the requirements of our proposed energy profiling system. These adaptations involved considerable modifications and tailoring to the functions and classes of the libraries, allowing for their smooth incorporation into our codebase. This collaborative strategy allowed us to effectively execute the essential components of the DRES generation profiling system, including data synchronisation, hyper-parameter optimisation, model training and evaluations.

4.2 Evaluation

4.2.1 ACF and PACF analysis

As a part of our pre-processing software component presented in Section 4.1.2.1, we utilize ACF and PACF analysis to test the correlation structure of the generated power measurements. Fig. 4.3 presents the result of ACF analysis. It can be observed that there is a high positive correlation with the lags outside of the 95% confidence interval.

We observe a high inter-correlation among the historical components of the generated power measurements from the ACF plot. In Fig. 4.3, each point on

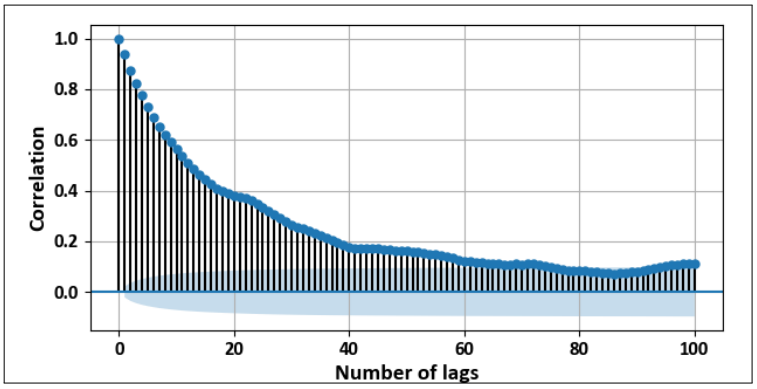

Figure 4.3: ACF of the generated power.

the ACF plot represents the correlation between the power generated at a specific time point and the power at a certain number of time steps back. The significant inter-correlation observed in this figure indicates that the power measurements are highly correlated at various lags. This can lead to unreliable statistical inferences due to multi-collinearity. Multi-collinearity occurs when independent variables in a statistical model are highly correlated, making it difficult to discern their individual effects.

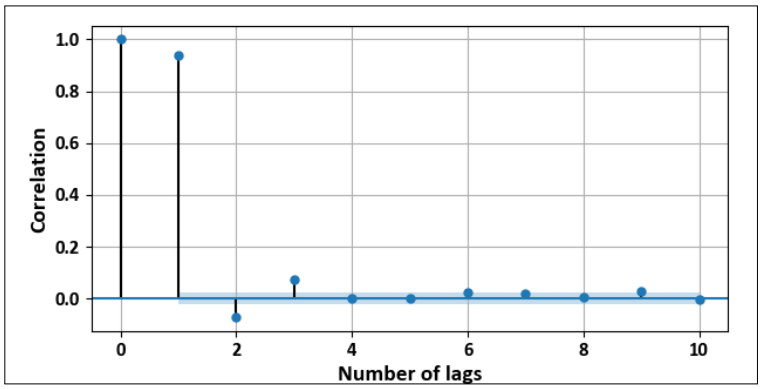

Figure 4.4: PACF of the generated power.

Therefore, we utilize PACF plots to retain only relevant lags, as opposed to the complete ACF plot, and to remove those that yield indirect correlations. The PACF plot is a useful tool for identifying direct correlations between the current observation and past observations, while controlling for indirect correlations through intervening time steps. In other words, it helps us isolate the unique influence of each lag on the current measurement, excluding the influence mediated through other lags. In Fig. 4.4, the PACF plot shows that lag_1 has the highest positive correlation after it first

intersects the confidence interval. Consequently, we use the values of lag_1 from the generated power as a feature input for the learning techniques in this study.

4.2.2 SCADA-based wind power modelling

As discussed in Section 4.1.3, our evaluation methodology firstly targets to compose a ground truth profiling model using SCADA-based measurements from the La Haute Borne wind farm. Hence, a total of 27 year-wide SCADA-based features were initially scrutinised by the feature selection component within the DRES profiling system presented in Section 4.1.2. The iterative feature selection process within the DRES profiling system has demonstrated that the RFE technique produced the best set with a total of 13 SCADA-based features for SVR, 6 features for KNNR and XGBoostR, 10 for UFS under the MLPNN models. In general, the filtered set of features is composed by a range of mechanical (e.g., pitch angle and generator converter speed), power (i.e., apparent power and the lag_1 feature) and weather (i.e., wind speed) features. Hence, these machine learning techniques covered all exogenous as well as intrinsic factors related to the wind-turbines behaviour in terms of power generation.

Under the combination of the selected features with the various machine learning-based regression components of the DRES profiling system, we have witnessed improved regression models in all of the machine learning-based algorithms. As illustrated by Fig. 4.5, the designed feature selection schemes positively impacts the performance of KNNR, SVR, XGBoostR and MLPNN, by reducing their MAE errors. Similar trends are also observed for the MSE. Overall, the SVR model under the RFE-based feature selection produced an extremely low MAE and MSE; MAE=0.000326 kW and MSE ≈ 0 kW^2.

4.2.3 SCADA-agnostic wind power modelling

Following the same pre-processing, normalization and feature selection performed within the DRES profiling system (as with the SCADA-based profiling), we have produced regression models using third-party weather features. The feature selection process identified 37 features from the three third-party data providers including measurements such as output temperature, pressure and wind direction, gust and speed. Subsequently these features were utilised within the core learning process for the XGBoostR, and 9 out of the 42 were employed within the SVM, KNNR and MLPNN models.

We observe that the prediction results for wind power regression based on the freely available third-party weather features varied slightly from the SCADA-based profiling. Nonetheless, the conducted experiments indicate no major difference in the obvious pattern for the estimated power curves as depicted in Fig. 4.6. Moreover,

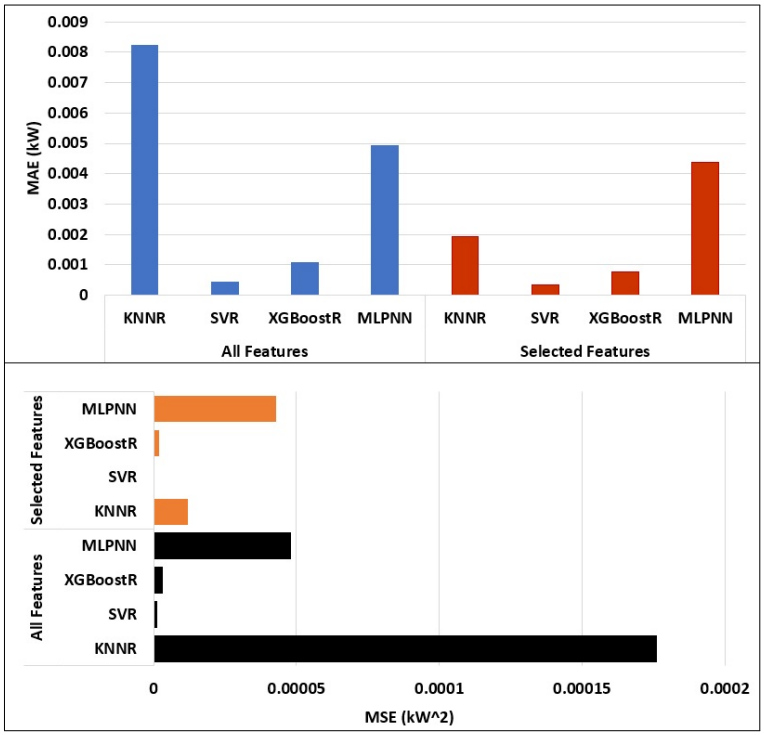

Figure 4.5: Errors between the actual and predicted power values based on SCADA measurements.

the performance analysis of the resulted machine learning-based regression models in relation to the MAE and MSE respectively shows that SVR outperforms the rest of formulations.

As evident, the SVR technique has a minimum MAE and MSE, where MAE is 0.003595 kW and MSE ≈ 0.0 kW^2. Meanwhile, the MAE for KNNR, XGBoostR and MLPNN are 0.027937, 0.004106 and 0.008734 kW, and MSE are 0.001533, 0.000038 and 0.000148 kW^2 respectively. Hence, the error performance shows a slightly higher error rate than SCADA-based but arguably to be of minimal importance for large-scale accounting and optimisation processes as required by the main grid. In parallel, under the scenario of a windfarm owner or third-party company with no access to SCADA measurements, we highlight that the approximate generation and potentially financial forecasting is not necessarily affected on a macroscopic scale. In addition, the actual SCADA-agnostic estimation is of minimal financial cost in comparison with a subscription-based SCADA-based approach as it usually happens.

As depicted in Fig. 4.7, the computational cost for producing a reasonable regression model is far smaller using a SCADA-agnostic approach in comparison to

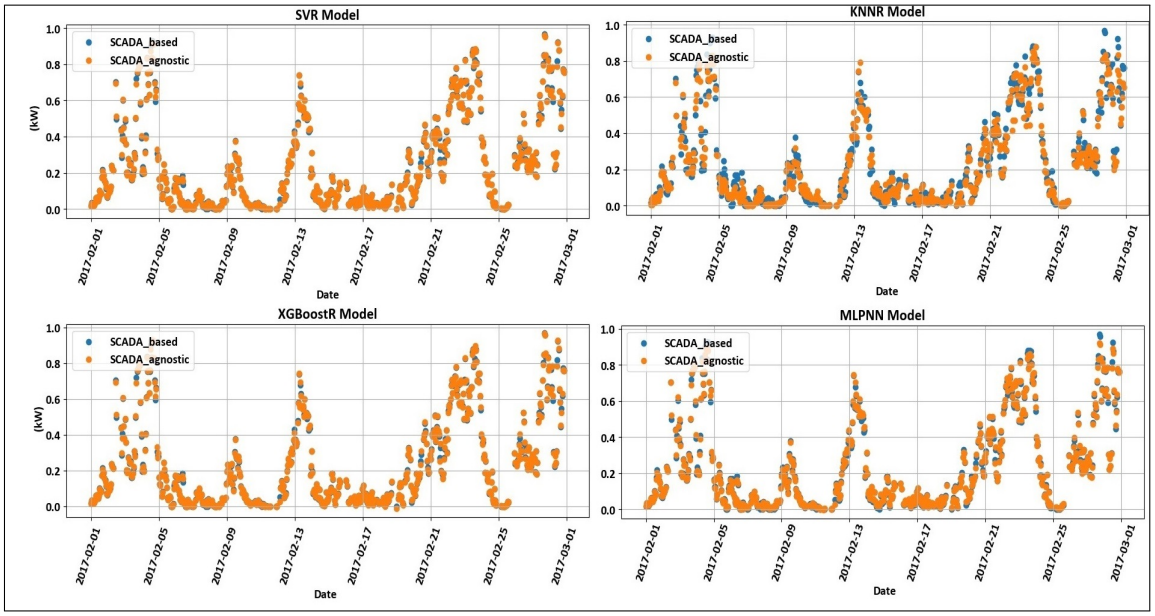

Figure 4.6: SCADA-based and SCADA-agnostic power curve. The y-axis represents the generated power in kW.

SCADA-based[6] approach. We also witness that the MLPNN model can act as a good approach for real-time use, however with some minimal trade-off with respect to their error rate performance. For long-term estimation processes, we observe that the XGBoostR alongside the SVR formulation promotes slightly more accurate SCADA-agnostic wind power profiling.

In general, the simplicity of utilising just three freely available features in comparison to expensive SCADA-based monitoring and measurement components could effectively pave the path towards new directions on real-time and low-cost DRES power profiling.

4.3 Summary

The increasing utilisation of DRES in the modern smart grid engages a complex energy trading model with vague policies in terms of hardware and software ownership in DRES deployments. Hence, it is not uncommon for independent DRES deployment owners to not have a complete control or access of their installations through SCADA systems managed by third-party providers or main grid operators. In this work,

[6]On a 64-bit Windows operating system with Intel Core i7 (7*th* Gen) CPU with 2.70 GHz clock cycle and 12 GB RAM.

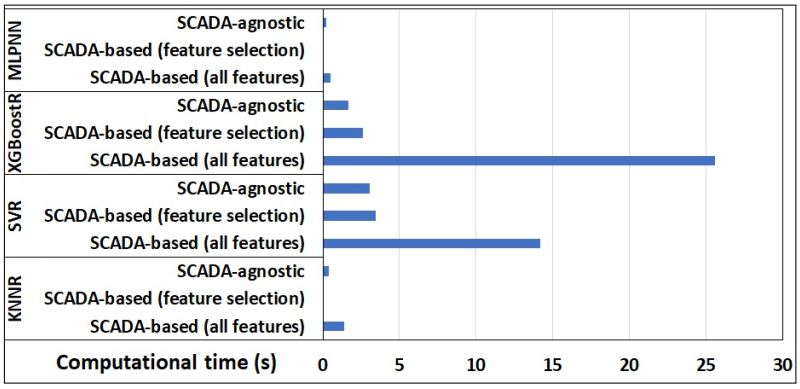

Figure 4.7: Computational time comparison.

we propose a SCADA-agnostic DRES profiling system and exhibit its applicability on a proof-of-concept study over a real wind turbine installation. We demonstrate that by simply utilising freely available third-party weather data with available regression models, we can reasonably match up to a great scale the regression accuracy performance of models utilising SCADA measurements. Moreover, the proposed SCADA-agnostic profiling is achieved with a minimal set of weather features in contrast to the SCADA-based approach and under a lower computational cost. Thus, paving the path towards independent and cost-efficient power generation profiling serving a range of envisaged smart grid applications such as virtual power plant design and malicious actor detection.

Chapter 5

Predictive Energy Theft Detection for Distributed Renewable Energy Sources

By virtue of global climate challenges, we witness a drastic shift by regulators and grid operators towards the full integration of distributed renewable energy sources (DRES) within modern smart grids with intriguing applications (e.g., virtual power plants). In fact, a number of developed and developing nations target 100% of energy generation to be resulted by DRES by 2040 (e.g., Sweden) and a 32% proportion to be achieved on average in the EU by 2030 (Vaughan, 2018). Nonetheless, the practical operation of such deployments entails a number of cybersecurity challenges that also transform the way in which energy theft could be manifested. Energy theft has been a traditional challenge, however, the refinement of the grid's business model and the relatively recent interconnection of DRES deployments with the main grid has enabled the composition of energy theft (Krishna, Gunter, and Sanders, 2018).

Numerous energy theft events are reported daily on a global scale affecting a range of operational factors for our society including safety and economy. For instance, non-technical losses caused by energy theft amount to £1.4B per annum in Brazil and for a single energy provider in Canada such thefts cause an average annual loss of 850 GWh converted as £55M of financial loss (Raggi et al., 2020). Evidently, both energy and monetary losses from energy theft are of paramount and timely importance, with direct implications to the general public's well-being as well as economy. Furthermore, the continuous and evolving manifestation of such events justifies the fact that current theft detection schemes employed by energy providers are inadequate.

As energy theft attacks increase, a momentum on the development of data-driven detection has been observed within the wider research community. However, a significantly small portion of detection solutions such as in (Yuan, M.-g. Shi, and

Sun, 2015; Krishna, Gunter, and Sanders, 2018; Mahmoud et al., 2020) considers the manifestation of energy theft from individual DRES owners. Moreover, the dependence on data that can be tampered with by adversaries (Krishna, Gunter, and Sanders, 2018; Yuan, M.-g. Shi, and Sun, 2015) or unavailable data in terms of supervisory control and data acquisition (SCADA) (Mahmoud et al., 2020) further restricts the reliability of these approaches in practical scenarios. In parallel, the aforementioned dependence on aggregated consumption SCADA measurements is unable to capture the intrinsic environmental dynamics such as to relate generation values with the actual weather conditions in a given DRES deployment, as we discussed in Chapter 4.

Therefore, in this work we take a practical approach by firstly proposing a generalised DRES-based adversary model and by secondly providing a predictive data-driven detection solution considering weather dynamics. This predictive energy theft detection approach is also an instantiation of our coherent framework proposed in Chapter 3. Utilizing the MMC dimension of the proposed theoretical framework, our predictive detection approach collects generation measurements from large-scale wind turbine and solar panel installations deployed across the infrastructure dimension and managed by DRES owners within the end-user dimension. The collected data are then utilised within the theft profiling sub-process of the proposed theoretical framework to propose adversaries who intend to launch an energy theft attack against DRES installations. The output of this round of theft profiling is then employed in the detector construction and theft detection sub-processes within the operation dimension of the proposed theoretical framework.

In general, the contribution of this work is two-fold by providing:

1. A formalised approach on describing DRES-based adversaries with the objective of energy theft. We demonstrate the efficacy in which DRES owners (i.e., prosumers) can take advantage of the current business model and gain financial benefits.

2. A novel, low-cost and generic energy theft detection framework comprising of two algorithms; i) a SCADA-agnostic DRES profiling method operating purely on third-party and widely available weather measurements and ii) a classification scheme relying on DRES profiling and able to classify theft detection events. Evidently, the synergy of the two components enables adaptive and highly accurate detection with low computational overheads and aiding significantly on reducing monetary loss.

The rest of this chapter is structured as follows: Section 5.1 provides a description of the system and Section 5.2 presents the adversary model for DRES-based energy theft attacks. Section 5.3 describes the methodology underpinning the proposed

detection framework, while Section 5.4 discusses the datasets used within this work. Section 5.5 depicts the evaluation methodology followed within our experimentation and Section 5.6 discusses the results obtained. Finally, Section 5.7 concludes and summarises this chapter.

5.1 System description

We consider an end-to-end energy system consisting of a single Transmission System Operator (TSO) connected with one or more Distribution System Operators (DSOs) consisting of nodes equipped with smart control, management, monitoring and metering technologies. The TSO is abstracted to a set of supply nodes R including DRES deployments and a set of high-voltage transmission buses denoted as Q. It is assumed that one or more DSOs of the set $P = \{p_1, .., p_n\}$ interact with the TSO in discrete time intervals and the energy supplied from the TSO to a given DSO on a discrete time interval is denoted as the function $Es(t)$. Energy transmission and distribution is achieved via bidirectional power and data communication flows through corresponding power system and communication control and management components (e.g., actuators, SCADA).

Each DSO in P is defined by a total number of nodes N and a set M of medium/low-voltage distribution buses. Nodes are categorised into A supply nodes and B demand nodes which we refer to as prosumers and consumers respectively, where $A \subset N$ and $B \subset N$. The energy produced by a single prosumer of A in the i^{th} DSO at a given discrete time interval t is mapped as the function $Er(t)$ whereas the energy consumed by a single consumer over a time period t in the i^{th} DSO is represented by $Ec(t)$. A prosumer is assumed to be an individual or a group of individuals owning and managing a DRES deployment (e.g., domestic solar panels) and can act both as a consumer and a supplier of energy back to the DSO.

As discussed in (Z. Zheng et al., 2018; Bihl and Hajjar, 2017), energy theft events cause energy losses that can be described as the difference between the generated energy and the energy consumed under normal conditions. Thus, we express the cumulative energy loss experienced for a single DSO in time t as:

$$L = \Delta Es(t) + \Delta \sum_{a=1}^{|A|} Er_a(t) - \Delta \sum_{b=1}^{|B|} Ec_b(t) + \sum_{m=1}^{|M|} TL_m(t) \tag{5.1}$$

where Δ is the discrepancy including the scaling discrepancy in meter readings for reported and actual measurements as caused by a single or more theft events at time t and TL refers to technical losses occurred due to physical constraints on transmission lines. Consequently, from a TSO perspective the total energy loss in

time t is expressed as:

$$L_{TSO} = \sum_{i=1}^{|P|} L_i(t) \tag{5.2}$$

where P is the total number of DSOs connected to the TSO.

5.2 Adversary model

The adversary model has an explicit focus on energy theft initiated by generation meters installed on DRES deployments and managed by prosumers. Thus, the primary assumption is that prosumers tamper generation meters and report erroneously back to their corresponding DSO. In order to reduce the complexity invoked within our adversary model, we consider the DSO meters interacting with edge DRES deployments to be secure. Thus, we rule out any discrepancy in the measurement function for energy supplied from a TSO to a DSO having $\Delta Es(t) = 0$.

In addition, we assume that smart-meters strictly reporting energy consumption by a given consumer to the DSO are not tampered. Therefore, discrepancies on the consumption reporting by all consumers in a given DSO complies with: $\Delta \sum_{b=1}^{|B|} Ec_b(t) = 0$.

As mentioned, we particularly focus on tampering conducted on individual meters reporting energy generation for a given DRES deployment. Hence, we deduce that: $\Delta \sum_{a=1}^{|A|} Er_a(t) \geq 0$.

Given the above assumptions we re-express the DSO energy loss as:

$$L = \Delta \sum_{a=1}^{|A|} Er_a(t) + \sum_{m=1}^{|M|} TL_m(t) \tag{5.3}$$

Based on Equation (5.2), which represents the total energy loss of a TSO, and Equation (5.3), which represents the non-technical energy loss caused by discrepancies DRES in a DSO, the energy loss for the TSO can be approximated as follows:

$$L_{TSO} = \Delta \Big\{ \sum_{i=1}^{|P|} \sum_{a=1}^{|A|} Er_{i,a}(t) \Big\} + \sum_{i=1}^{|P|} \sum_{m=1}^{|M|} TL_{i,m}(t) \tag{5.4}$$

In order to cover a spectrum of tampering behaviour by a prosumer, we define four types of theft functions. All four functions mimic practical fraudulent patterns in terms of reporting erroneous generated energy back to the DSO. Many possibilities of such theft functions exist and the herein models are distilled by observations in

literature (Mahmoud et al., 2020; Yip, Wong, et al., 2017b; K. Zheng et al., 2018; Shaaban et al., 2021). Through our work, we emulate attackers that attempt to create manipulated reports either by retaining original curve fluctuations and features or by generating new patterns. From a modeling perspective, these are variables that partially or completely amplify the reported energy timeseries signal as we show next.

1. Total Scaling Theft:
$$\Delta Er(t) = \eta(t) Er(t) \tag{5.5}$$
where $\eta(t) \in \mathbb{R}$ and $1 < \eta(t) < \infty$.

2. Partial Scaling Theft:
$$\Delta Er(t) = \begin{cases} Er(t), & Er(t) \geq \beta \\ \beta, & Er(t) < \beta \end{cases} \tag{5.6}$$
where $\beta \in \mathbb{R}$, $\beta > \min\big(Er(1), Er(2), ..., Er(D)\big)$ and D is time period equaling to one day.

3. Off-Peak Theft:
$$\Delta E(t) = \gamma E(t) \tag{5.7}$$
where
$$\gamma = \begin{cases} \eta, & t \in [t_{start}, t_{end}] \\ 1, & otherwise \end{cases}$$
where $[t_{start}, t_{end}]$ is the off-peak period, that is the peak operating weather conditions for DRES.

4. Replay Theft:
$$\Delta Er(t) = \max\big(Er(1), Er(2), ..., Er(D)\big) \tag{5.8}$$
where D is a discrete time period that equals to one day (24 hours).

In more detail, the total scaling theft in Equation 5.5 considers the scenario in which the aggregated generation measurements on time t are tampered by an attacker. Tampering is based on an arbitrary percentage denoted by η, which is adjustable (i.e., random rate percentage). For instance, the attacker reports 150% of the actual measurements when $\eta = 1.5$. The partial scaling theft scenario in Equation 5.6 considers the case where an adversarial prosumer tampers generation measurements

whilst a particular threshold is met. Hence, the prosumer sets a minimum reporting value (i.e., β) for the DRES-based generation measurements sent to the DSO.

We also consider the case in which theft could be temporally sporadic. Thus having discontinuous reporting of erroneous generation measurements during an off-peak period that relates with the peak weather conditions in which the DRES operates. For instance, the fraudulent prosumer reports 40% more power for a given time period than what was actually generated during the off-peak solar radiation of that period. Therefore, only measurements generated during the off-peak operating conditions of DRES are scaled as shown in Equation 5.7. Finally, in the replay attack, the attacker only reports the highest actual generation for the whole time duration T.

5.3 Methodology

The data flow underpinning the proposed framework is depicted by the flowchart in Fig. 5.1. Building upon the work in Chapter 4 where concrete profiling and prediction of energy generation using purely geo-located weather features is achieved, we profile DRES installations using third-party and widely available weather measurements. The incentive behind this approach is to remove the full dependency on SCADA-based measurements. Hence, we align with the realistic scenario where no available measurements gathered by locally placed sensors or the DRES SCADA systems exist. As depicted, the SCADA-agnostic profiling component works in synergy with a classification component that considers reported DRES energy generation measurements as seen at the DSO level. Thus, tailor theft detection over individual DRES installations and back-track potential fraudulent prosumers.

5.3.1 SCADA-agnostic DRES energy profiling

As shown in Fig. 5.1, the implemented DRES energy profiling component accepts third-party weather measurements and it first employs an automated pre-processing procedure. Following a series of data-oriented tasks dealing with noisy and incomplete measurements, the profiling component conducts an automated feature selection process in which the most suitable statistical features are chosen to compose a DRES energy generation profile. The trained model based on the selected subset of features of the source DRES is then used for all DRESs to profile their generated energy measurements. Details of all the processes involved in the aforementioned description in terms of profiling are discussed next.

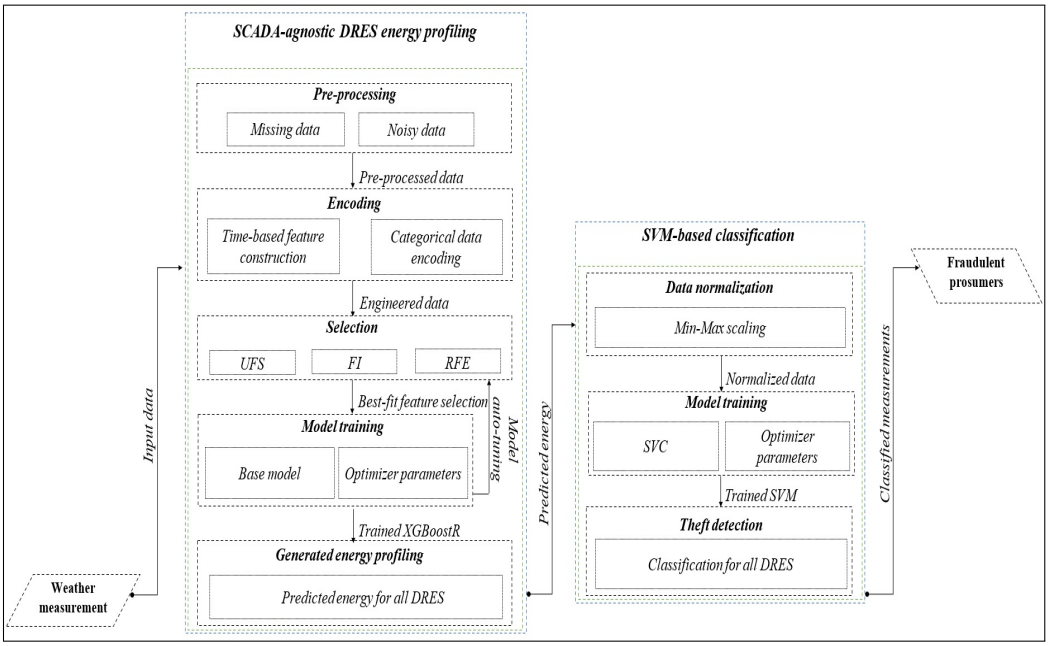

Figure 5.1: Data flows defining the proposed energy theft detection framework.

5.3.1.1 Data pre-processing

Within the pre-processing stage we filter our raw measurements by removing all the possible missing, duplicated and inconsistent samples. Usually, third-party measurements are largely inconsistent (out-of-range) due to various factors ranging from environmental sensor reading and reporting errors as well as REST-API pull failures. In this context, we include common data sanitisation and normalisation approaches within an automated pipeline in our prototype.

5.3.1.2 Data encoding

The encoding process enables granular representations from aggregated third-party time series. Given that aggregated measurements contain categorical time series, we encode them into numerical data via a binary encoder. In more detail, we assign an integer value to each unique category of the original categorical vector. Subsequently, for each integer-encoded category we generate a binary vector and based on the majority of bit coding, we generate an additional measurement vector. Finally, we construct temporal views of our categorical measurements using the measurement timestamps having hourly, daily and monthly observations mapped with the aforementioned binary encoded measurement vector.

5.3.1.3 Data selection

The developed data selection component is underpinned by an automated feature selection mechanism such as to identify an appropriate set of features from the encoded timeseries described earlier. As shown in Fig. 5.1, our component works in coordination with the model trained on the source DRES component to obtain the optimal feature set, producing the best-fit prediction model of the DRES generated power. In detail, the selection process utilises i) univariate feature selection (UFS), ii) ranking-based feature importance (FI) and iii) wrapper-based recursive feature elimination (RFE). The reason of using three feature selection algorithms is to ensure that we compile the best combination of meta-features. The importance of each of the features is compared and chosen based on the F-score and the Pearson correlation coefficient as well as a random forest estimator.

5.3.1.4 Model training

The base DRES profiling model strongly depends on the aforementioned feature selection process. Most importantly, it is dynamically updated whilst new and improved data feature combinations are provided by the selection process. Through a repetitive feedback mechanism and the continuous update of a boosting regressor we achieve an adaptive DRES energy generation base profile.

Due to its scalability in several measurement distributions, we utilise the XGBoostR algorithm proposed by Chen and Guestrin (T. Chen and Guestrin, 2016). This algorithm is composed of a classification and regression tree (CART) ensemble model using K additive functions. Thus, enabling prediction of the generated power measurements of the DRES installations at DSO. The proposed prototype minimises the regularised objective of the XGBoostR model as defined in (T. Chen and Guestrin, 2016):

$$\mathcal{L} = \sum_{t=1}^{|T|} l\big(y(t), \hat{y}(t)\big) + \sum_{k=1}^{|K|} \Omega(f_k) \tag{5.9}$$

where $y(t)$ and $\hat{y}(t)$ denote the actual and predicted power measurements at time t respectively, l is the loss function measuring the difference between $\hat{y}(t)$ and $y(t)$, f_k refers to the k^{th} tree structure and Ω denotes the model complexity for avoiding over-fitting which can be expressed as:

$$\Omega(f) = \zeta J + \frac{1}{2}\lambda \mid\mid w \mid\mid^2 \tag{5.10}$$

where J is the number of leaves in the tree, w is the leaf weight, and ζ and λ are constants controlling the regularisation degree.

Since the XGBoostR model is trained in an additive fashion, Equation 5.9 at the j^{th} iteration can be expressed as:

$$\mathcal{L}^{(j)} = \sum_{t=1}^{|T|} l\Big(y(t), \hat{y}(t)^{(j-1)} + f_j\big(x(t)\big)\Big) + \Omega(f_j) \tag{5.11}$$

where $x(t)$ is the input vector at t and $\hat{y}(t)^j$ represents the power measurement prediction of the t^{th} observation at the j^{th} iteration.

Following (T. Chen and Guestrin, 2016) within our implementation, we used second-order approximation to optimise the objective, which can be simplified as follows:

$$\mathcal{L}^{(j)} = \sum_{t=1}^{|T|} \Big(g(t) f_j\big(x(t)\big) + \frac{1}{2} h(t) f_j^2\big(x(t)\big)\Big) + \Omega(f_j) \tag{5.12}$$

where

$$g(t) = \partial_{\hat{y}^{(j-1)}} l\big(y(t), \hat{y}(t)^{(j-1)}\big) \tag{5.13}$$

and

$$h(t) = \partial^2_{\hat{y}^{(j-1)}} l\big(y(t), \hat{y}(t)^{(j-1)}\big) \tag{5.14}$$

The optimal hyper-parameters for the employed XGBoostR model training process were obtained using a grid search-based cross-validation technique returning the appropriate values with the lowest prediction error.

5.3.1.5 Generated energy profiling

The generated energy measurements of all DRES at the DSO are predicted using the trained XGBoostR as follows:

$$\hat{y}(t) = \sum_{k=1}^{|K|} f_k\big(x(t)\big) \tag{5.15}$$

where f_k denotes the k^{th} tree structure.

Accordingly and due to the cumulative nature of the utilised XGBoostR, the predicted power measurements of the source DRES at step j can be calculated as follows:

$$\hat{y}(t) = \hat{y}(t)^{(j-1)} + f_j\big(x(t)\big) \tag{5.16}$$

As discussed next, the resulted features are utilised within the SVM-based classification process in order to detect fraudulent prosumers.

5.3.2 SVM-based classification

Due to its optimal performance in terms of memory efficiency (Jindal, Dua, et al., 2016), a supervised SVM classifier is trained based on the predicted energy measurements calculated using Equation (5.15) such as to provide a binary prediction class for each of the prosumers in a given DSO (i.e., fraudulent or not).

As shown in Fig 5.1, the first process within the implemented SVM-based classification prototype deals with normalisation of the DRES profiling output, which is subsequently used as input to the DSO's SVM training model. As explained next, the classification phase is decomposed into specific processes in order to ensure unbiased detection of fraudulent prosumers.

5.3.2.1 Data normalisation

Prior the training and classification stage we employ a min-max normalisation technique to reconstruct the processed measurements in the range of $[0, 1]$. Normalisation is a crucial component within any statistical representation process and particularly in our case we achieve to ensure testing and not neglecting extremely small measurement values.

5.3.2.2 Model training

Following data normalisation, a model classifier is resulted by processing the training set to detect fraudulent prosumers. Thus, the predicted energy measurements from Equation (5.15) are the input to a trained SVM model in such a way to accommodate an optimal decision boundary for classifying DSO prosumers. The optimal SVM hyperplane boundaries can be obtained by solving the following soft optimisation problem:

$$min\Big(\frac{1}{2} \mid\mid w \mid\mid^2 + C \sum_{t=1}^{|T|} \xi(t)\Big) \tag{5.17}$$

where w denotes the weight vector, C denotes the regularisation parameter used to quantify the trade-off between the model's complexity and the classification error. Also, ξ represents a slack variable.

In order to select the most appropriate hyper-parameter values with the highest training accuracy for the SVM model we employ a synergetic use of grid search and cross-validation algorithms.

5.3.2.3 Theft detection

Once the SVM-based training model is achieved, the binary classification of DSO prosumers to either being legit or fraudulent is conducted.

The decision boundary function in our proposed implementation is defined as:

$$f(x) = \sum_{t=1}^{|S|} \big(\alpha(t) - \beta(t)\big) K\big(x(t), y(t)\big) + b \tag{5.18}$$

where S is the set of support vectors, $x(t) \in S$ is the support vector, $y(t)$ is the assessed power measurement, K is the kernel function. The $\alpha(t)$ and $\beta(t)$ variables are the Lagrange multipliers and b is the regularisation parameter.

Due to the fact that all energy measurements have a non-linear distribution, we employ a radial basis function (RBF) kernel defined as:

$$K\big(x(t), y(t)\big) = exp\Big(- \gamma \mid\mid x(t) - y(t) \mid\mid^2 \Big) \tag{5.19}$$

where γ is the kernel function parameter.

5.4 Dataset description

Our evaluation is based on real measurements gathered by wind turbine and solar panel installations in Australia and France. In particular, we utilise a 10-fold cross-validation by the mean technique on a dataset acquired from the La Haute Borne wind farm located in Meuse, France [1] and a solar power dataset captured at the Ausgrid power network located in Sydney, Australia [2].

Table 5.1 depicts a summary of the aforementioned datasets. As shown, the Engie wind datasets represents the daily generated power measurements captured at a real installation of 4 wind turbines for a duration of 11 months in 2017. In addition, the Ausgrid solar data provides daily measurements captured for a period of 11 months from rooftop solar panel installations.

Table 5.1: Datasets overview.

Dataset	Time Window		Location		DRES Capacity
	Start	End	Longitude	Latitude	
Engie Wind	Jan 2017	Dec 2017	5.6013 E	48.4503 N	2050 kW
Ausgrid Solar	Jul 2012	Jun 2013	151.2093 E	33.8688 S	1 kW

[1] Explore – ENGIE France Renewable Energy Open Data, Available: `https://opendata-renewables.engie.com/pages/home/`

[2] Explore – Ausgrid Solar Home Electricity Data, Available:`https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data`

Within Table 5.1, we highlight longitude and latitude values since they were critical for mining weather and environmental information explicit to those areas. Hence, we extracted available measurements such as output temperature, wind speed, humidity and pressure and assessed their ground truth by cross-validating across multiple third-party and freely available APIs. In order to do that, we collected data from the Dark Sky API[3], Weather Online[4] and Open Weather API[5] aligning with the same observational period as that of the generation measurements at the Engie and Asugrid installations. Complementary to the aforementioned, we acquired additional output temperature and wind-related measurements from Weathernews[6] as observed by the Nancy-Ochey weather station, which is geographically adjacent to the La Haute Borne wind farm. In total, our third-party weather data, i.e., the obtained weather measurements from freely available APIs, comprised of 52 weather and environmental measurements, including numerical and categorical readings such as wind measurements, humidity, pressure and cloud cover within hourly sampling bins. For our evaluation group, measurements are considered seasonally (i.e., summer, autumn, spring, and winter).

5.5 Evaluation Methodology

The evaluation methodology employed within this work aims at determining the suitability of the integrated data-driven theft detection solution over diverse DRES deployments. The main focus was placed on quantifying the detection performance and also relating it with the corresponding computational costs. Moreover, we conduct a monetary meta-analysis assessing the potential impact of the various synthetic thefts as well as the theft detection gains from the DSO perspective.

5.5.1 Detection performance

A thorough analysis was conducted such as to evaluate the detection performed using the synergy of the SCADA-agnostic DRES energy profiling and the SVM classifier discussed in Section 5.1. The first phase consists of the SVM classifier training with input from the SCADA-agnostic DRES power profiling output to compute final predicted power for all DRES installations. In parallel, an instance of the SVM component is trained based on the third-party weather data measurements. Within

[3]Explore – Dark Sky API, Available: `https://darksky.net/dev`

[4]Explore – World Weather Online API, Available: `https://www.worldweatheronline.com/developer/api/`

[5]Explore – Open Weather API, Available: `https://openweathermap.org/api`

[6]Explore – Weathernews, Available: `www.weathernews.fr`

our classification procedure, we distributed the dataset so that 70% of it is for training and 30% from each season as testing (Jindal, Dua, et al., 2016).

As already described in Section 5.1, we reach a binary detection decision (i.e., fraudulent or legit) through comparing the two outcomes of the aforementioned classification processes. Two classification errors and one computational cost metric were also utilised to assess the resulting classification models. The classification error metrics are accuracy (ACC) and binary classification area under the curve (BAUC) (related to receiver operating characteristic (ROC) curve), while we consider the time taken to obtain a decision as the computational cost. The definitions of which are provided as follows.

1. **ACC:** the ability to correctly differentiate fraudulent and legit measurements, defined as:
$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.20}$$
where TP, TN, FP and FN represent the true positives, true negatives, false positives and false negatives, respectively. TP is the number of measurements correctly identified as fraudulent, TN is the number of measurements correctly identified as legit, FP is the number of measurements incorrectly identified as fraudulent, and FN is the number of measurements incorrectly identified as legit.

2. **BAUC:** the degree of the capability of distinguishing between fraudulent and legit measurements and is defined as:
$$BAUC = \frac{\sum \text{Rank}^{(+)} - \frac{NF(1+NF)}{2}}{NF \times NL} \tag{5.21}$$
where $\sum \text{Rank}^{(+)}$ represents the sum of the ranks from fraudulent measurements, NF is the number of the fraudulent measurements, and NL is the number of legitimate measurements. Following (Z. Zheng et al., 2018; Hand and Till, 2001), we have arranged the measurements in ascending order based on the prediction of fraudulent measurements for ranking.

3. **Computation time complexity:** the time required by the SVM prototype to produce a DRES classification of a given DSO_i.

5.5.2 Theft scenarios

Due to the fact that the acquired datasets were the result of prosumers that volunteered to provide their data, we assume that all measurements were legit and no fraudulent behaviour is present. Hence, prosumers reported genuine generation

measurements therefore the original data are considered as the ground truth. As presented in Section 5.3, we inject synthetic anomalies in our datasets that conform to specific theft scenarios discussed in the literature. Hence, we employ our developed energy theft functions, i.e., (definitions 5.5), ((5.6)), (5.7) and (5.8) in order to compose a dataset consisting of both legitimate as well as fraudulent patterns. As depicted by Table 5.2, the conducted evaluation methodology considers varying theft proportions (i.e., fraudulent measurements) injected within the actual dataset across a given DSO. We stretch the scaling parameters for both total and partial thefts within particular boundaries to ensure that we create the most representative realistic scenarios. The mix theft scenario focuses on a randomly chosen subset of measurements to simulate one of the four main scenarios. As discussed in the literature, there exist many cases in which fraudulent prosumers might apply different theft scenarios over different time-periods to manipulate with their measurements (K. Zheng et al., 2018).

Table 5.2: Simulation parameters used in theft scenarios.

Dataset	Theft Scenario	Parameter
Engie Wind	Total Scaling Theft	$1.4 \leq \eta \leq 7$
	Partial Scaling Theft	50 kW $\leq \beta \leq$ 100 kW
	Off-Peak Theft	$rated\ wind\ speed = 15$ m/s
	Reply Theft	$D = 24$
Asugrid Solar	Total Scaling Theft	$1.4 \leq \alpha \leq 7$
	Partial Scaling Theft	0.005 kW $\leq \beta \leq$ 0.4 kW
	Off-Peak Theft	$11\,am \leq t \leq 3\,pm$
	Reply Theft	$D = 24$

5.6 Results

5.6.1 Theft detection performance

The results of the theft detection framework proposed in Section 5.3 are illustrated in this section, while considering the discussed evaluation methodology in Section 5.5 on the datasets specified in Section 5.4.

Using the SVM-based classification system proposed in Section 5.3.2, we witness that the SVM-based classifier trained on the energy profiling outperformed the classifier based on the third-party weather data for with regard to the ACC and BAUC scores in Engie wind data as shown in Figs. 5.2 and 5.3. In this case, total scaling theft results in 5.9% higher score when we use predicted wind energy output as a feature to train the SVM classifier, as compared to using the third-party weather data. In addition, the performance of the model based on third-party weather data significantly drops on the partial scaling theft scenario, with a margin of more than

13.16% compared to the first case when it was trained on the energy profiling output. Nonetheless, under the replay theft scenario, both classifiers achieved high ACC score of 97.2% and BAUC 99.6%. In the replay theft scenario, fraudulent prosumers over-report the maximum of the actual generation of the installed DRES. This theft behavior shows a steady and repetitive distribution throughout the fraudulent energy measurements, that can be unambiguously detected by the proposed SVM-based classification prototype in both cases, i.e., either it is trained using the third-party weather data, or energy profiling.

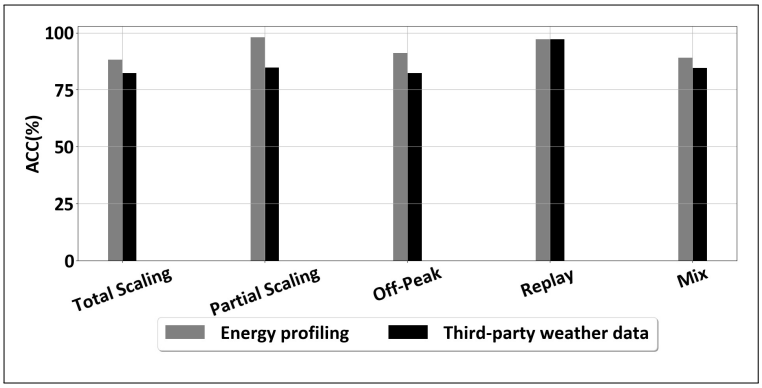

Figure 5.2: ACC values of the Engie wind power data.

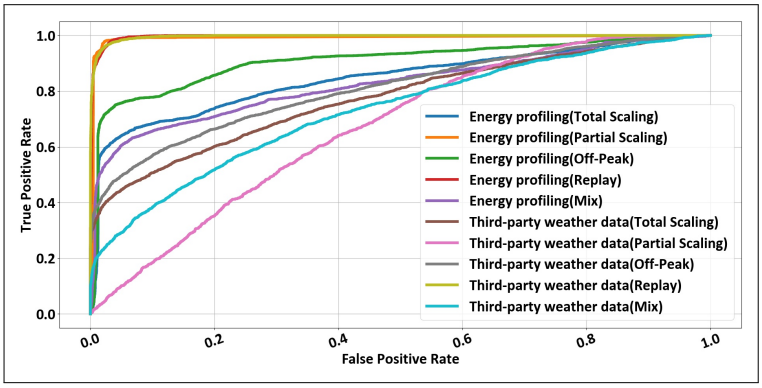

Figure 5.3: BAUC values of the Engie wind power data.

With respect to the Asugrid solar power measurements, the SVM-based classification prototype trained on the energy profiling output provided higher scores in detecting several theft scenarios, as shown in Figs. 5.4 and 5.5. As evident from these figures, in the case where the energy profiling output was used as an input feature, the SVM-based classification prototype obtained an ACC of 89.1% and an

BAUC score of 81.4% in detecting the mix theft. However, in detecting the same theft scenario, the SVM-based classification prototype trained on the third-party weather data maintained a lower score (more than 6%) than the one trained on energy profiling output by obtaining an ACC of 82.3% and an BAUC score of 74.2%. Similarly, to detect the total scaling theft, the SVM-based prototype trained on the energy profiling output provided more than 2% higher results than that trained on the third-party weather data. In detecting replay theft, both classifiers performed excellent scores on both ACC and BAUC.

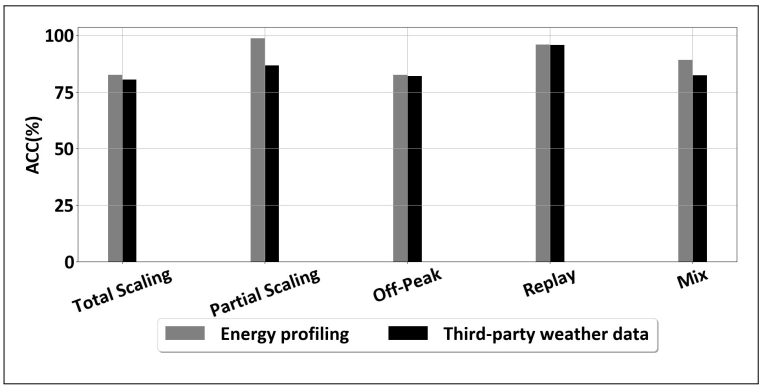

Figure 5.4: ACC values of the Asugrid solar power data.

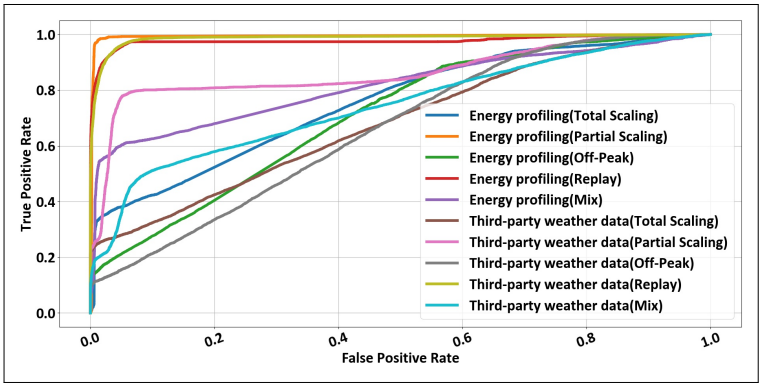

Figure 5.5: BAUC values of the Asugrid solar power data.

Overall, the SVM-based classification prototype trained on the energy profiling output maintained higher ACC and BAUC scores in both the wind and solar power measurements. Therefore, we can infer that the output of the SCADA-agnostic power profiling prototypes (i.e., the energy profiling for the DRES) has an important role to play in differentiating fraudulent prosumers. For more insight, Fig. 5.6

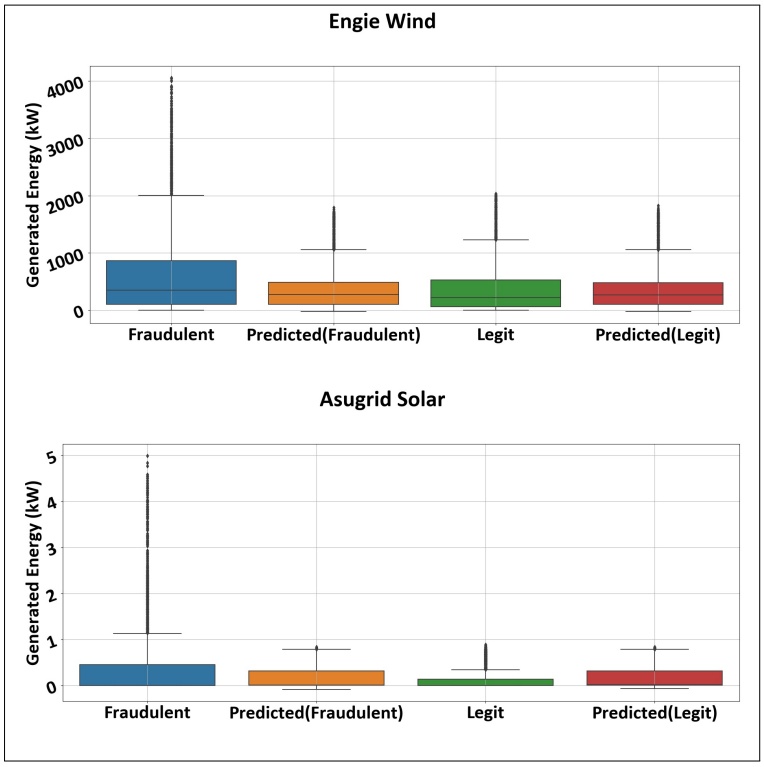

Figure 5.6: Predicted and actual power generation of legit and fraudulent prosumers in Engie and Ausgrid.

presents the boxplots of predicted power measurements and the actual generations for both legit and fraudulent prosumers. It is evident, that the energy measurements proportion range for legit prosumers was between 0 kW and the capacity of the DRES installations, thus 2050 kW for wind and 1 kW for solar respectively.

However, the energy measurement proportion of the fraudulent prosumer exceeded this range by the value of the manipulated green energy units. Moreover, it is demonstrated that the predicted power of the legit prosumers falls within 8% of actual generation, whereas a significant difference between the predicted and actual generated power can be observed for the fraudulent ones. Therefore, the predicted power measurements for DRES can be used as a useful feature for the proposed SVM-based classification prototype, where prosumers are classified either as fraudulent or legit based on their respective predicted energy measurements.

Fig. 5.7 depicts the results of the SVM-based classification prototype in both the proposed cases in terms of computational time in our evaluation methodology. This analysis was performed using a 64-bit Windows operating system with Intel Core i7 ($7th$ Gen) CPU with 12 GB RAM and 2.70 GHz clock cycle. The results

clearly indicate that the SVM-based classification prototype trained on the energy profiling prototype output operates on a relatively lower computational time than that on third-party weather data. The reason behind this is the high dimensionality of the aggregated third-party weather data, where the total number of selected measurements was more than that output of the energy profiling prototype, i.e., the predicted energy measurements. In the case of the third-party weather data, the process of classifying such high dimensional data requires computational complexity than that occurring in low dimensional spaces, where the SVM-based classifier only performs on the energy profiling output.

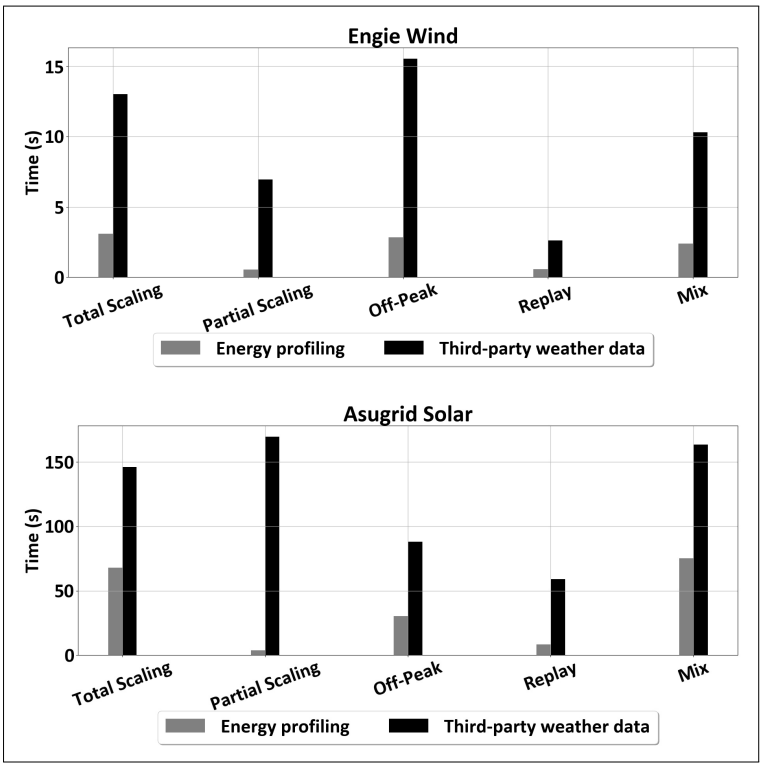

Figure 5.7: Computational time comparison.

5.6.2 Monetary analysis

The density of the amount of the energy loss caused by the theft attacks originally for the DSO in each month of the year is illustrated in Fig. 5.8. The amount of the energy losses in both wind and solar data can be obtained using Equation (5.3). Fig.5.9 presents the amount of the monetary cost caused by each individual theft scenario for the whole of a year. These monetary costs are estimated by multiplying the electricity price with the resulted total DSO energy loss in each season. The

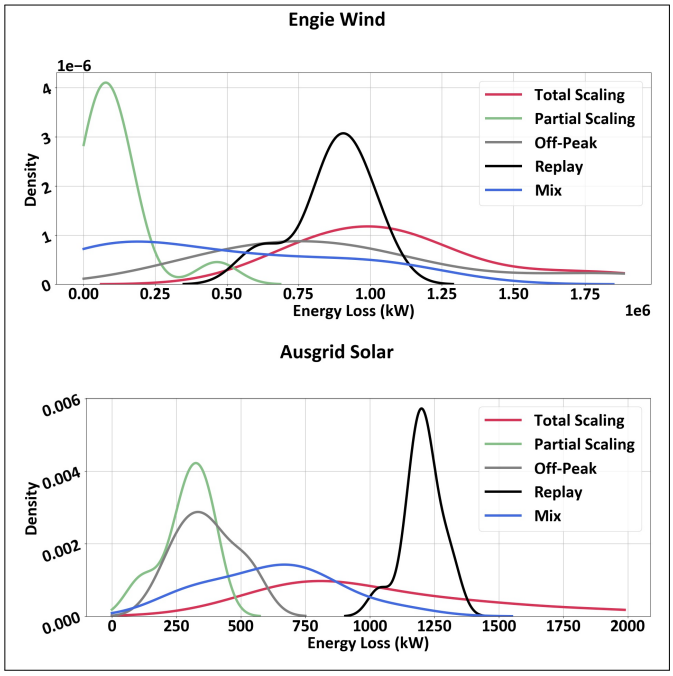

Figure 5.8: The density of the energy loss in wind and solar energy data.

France feed-in tariff (i.e., £7.40/kWh (Tazi and Bouzidi, 2020)) was applied to the wind energy dataset, while the Ausgrid feed-in tariff (i.e., £0.051/kWh (Krishna, Gunter, and Sanders, 2018)) was applied to the solar energy dataset.

These figures indicate that the monetary cost for the utility provider varied linearly with the amount of energy loss for both datasets. The spikes of the density curves in Fig.5.8 denote that for the wind measurement, the highest concentration of the highest energy loss was caused by the total theft scenario resulting in a monetary cost of 31% for the providers. The same behaviour was identified for solar measurements, where the highest concentration of the highest energy loss was caused by replay thefts, resulting in the largest amount of cost of 34.1% of the total annual monetary cost. As evident from these figures, the monetary costs can reach an incredible level when large-scale DRES are manipulated, especially for the replay or total scaling thefts. In such cases, the fraudulent prosumer engaged in theft activities manipulates the energy generation values measured by the generation meters endowed within the private DRES by increasing the number of green energy measurements that are reversed to the energy grid. Consequently, the energy losses increased by the discrepancy in this value, leading the utility provider to overcharge.

In order to save such monetary costs incurred by the providers, an accurate detection of the energy loss caused by energy thefts is required in the first place. Fig.5.10 presents the saved cost provided by our proposed detection framework. As

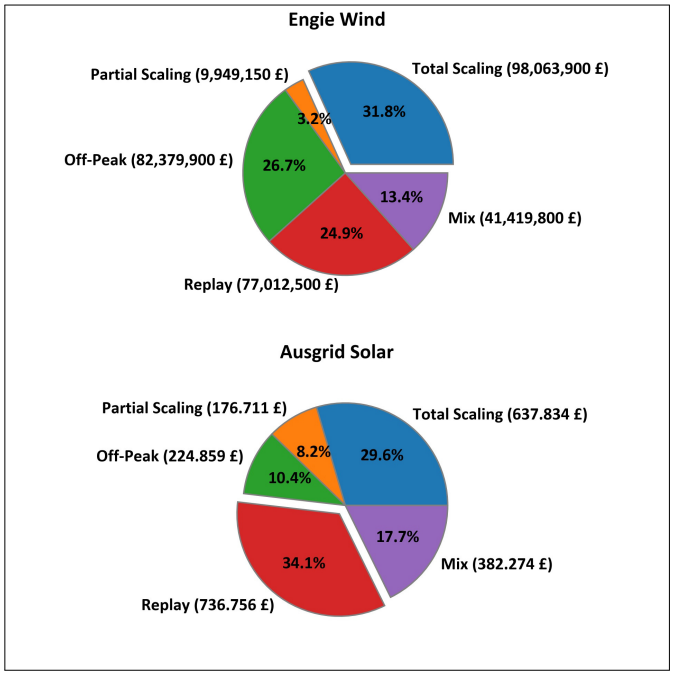

Figure 5.9: The amount of the monetary cost for the utility providers.

evident from this figure, for both wind and solar energy measurements, our framework saved between $82 - 99\%$ of the monetary cost through detection the energy loss. The median value of the cost that can be saved by the proposed detection framework of the total theft scenario in the wind energy measurement is £7, 808, 430, while in the solar energy measurements is £60.0395. The estimate of saved costs provided by the proposed framework is without including any additional hardware equipment since the proposed framework is completely data-driven, or utilizing additional measurements that are directly unavailable to the providers that are only aware of the DRES capacity.

5.7 Summary

Energy theft attacks pose a pressing issue that has resulted in enormous non-technical energy and monetary losses to energy providers at a global scale. The integration of DRES deployments in modern energy grids in conjunction with the widely adopted business model of demand-response have undoubtedly expanded the energy theft attack surface. Conventional energy theft detection schemes heavily rely on the assessment of spatiotemporal patterns from aggregated and commonly incomplete SCADA measurements without considering the intrinsic weather or environmental patterns related to a specific DRES deployment. Therefore, in this chapter we propose

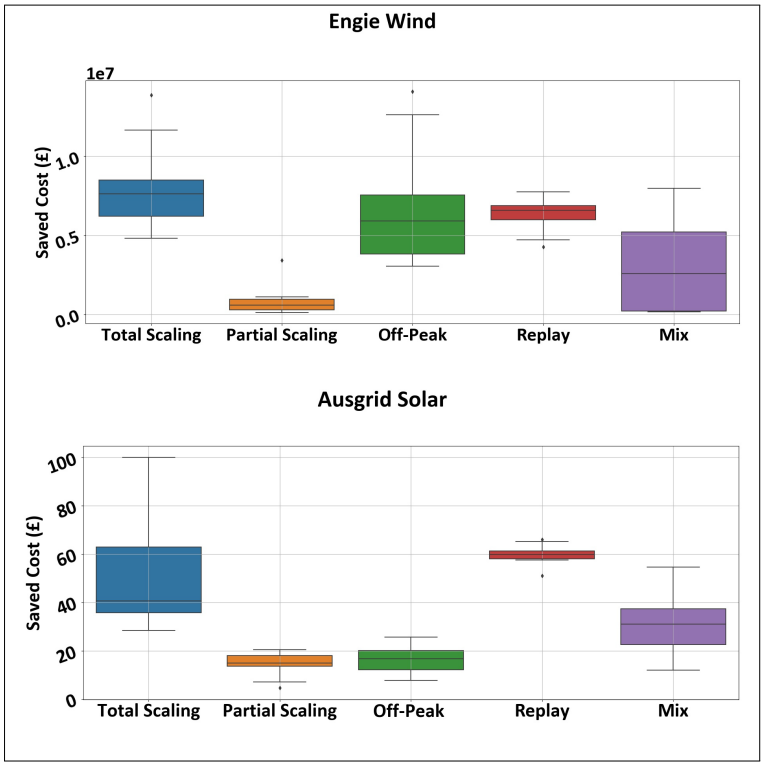

Figure 5.10: The saved cost by the proposed framework in wind and solar energy data.

a predictive SCADA-agnostic energy theft detection framework explicitly to DRES-based scenarios. We introduce a DRES-based theft attack model and further evaluate the performance of our framework by utilizing freely available third-party weather measurements over real solar and wind energy deployments in Australia and France respectively. Through our evaluations based on energy profiling model and third party weather data, we demonstrate that the proposed framework can detect fraudulent prosumers with high average accuracy with relatively low computational costs. Hence, placing it as a good and cost-effective candidate for future data-driven energy theft detection schemes.

Chapter 6

Adaptive Energy Theft Detection for Generation and Consumption Smart Meters

The modernisation of traditional power grids into smart grids through the demand response (DR) paradigm alongside the integration of distributed renewable energy sources (DRES) within grid optimisation practices is undoubtedly contributing towards the global net-zero initiative. A core property of modern power grids revolves around the adequate operation and optimisation of AMI as underpinned by networked smart meters. In the United Kingdom alone, the number of smart meter deployed by utility companies reached 28.8 million by the first quarter of 2022, with a full coverage projected by the end of 2025 (Gov.UK, 2022; Ofgem, 2022).

Given the diversity of hardware and software technologies entailed within smart meter integration and the lack of holistic grid-specific cybersecurity practices, we witness an evolving threat landscape in which energy theft activities are further enabled (Shaaban et al., 2021). Energy theft has been a problem since the very early coal-based power grids and its manifestation has changed dramatically with the introduction of networking technologies, as well as, advanced energy trading platforms. Estimates of monetary loss attributed to energy theft in the United Kingdom and the US have been put at $170 million and $6 billion, respectively, in the last few years (M. Wen et al., 2021). Evidently, the cybersecurity loopholes inherited by the interface of IoT technologies with AMI and legacy or bespoke industrial control system (ICS) in modern grids constitute the basis for various threat vectors in which consumers or prosumers could exploit primarily for monetary gain, as we discussed in Chapter 2.

By virtue of the direct relationship and impact of the DR paradigm with energy trading as translated into financial transactions, energy theft has received

a considerable level of attention by a number of studies. Nonetheless, the majority of studies was limited in scope due to their explicit focus on particular types of measurements or properties of the overall smart energy systems. Effectively, existing energy theft detection schemes rely on readings related to consumption (Z. Zheng et al., 2018; Yao et al., 2019; M. Wen et al., 2021) or focus on DRES energy generation measurements (Mahmoud et al., 2020; Shaaban et al., 2021).

Therefore, in this chapter, we aim to present an adaptive energy theft detection approach that considers both power generation and consumption measurements. The proposed method is capable of distinguishing theft-related events from the noisy data generated by misconfigured devices and, more importantly, it can self-optimise by utilising the properties of incoming measurement streams, without human intervention. The adaptive energy theft detection approach proposed in this chapter is another instantiation of our overarching framework described in Chapter 3. Hence, it employs the MMC dimension of the proposed theoretical framework and gathers data at various levels of aggregation in the infrastructure dimension. The collected data are then used within the operation dimension to extend the theft profiling sub-process of the proposed theoretical framework, with the aim of proposing a general adversary model applicable to stealthy energy theft and abnormalities resulting from legitimate consumption and generation measures. The output of this theft profiling round is then utilised to continue the detector construction, theft detection, and theft classification sub-processes within the operation dimension of the proposed theoretical framework.

In summary, we contribute by:

1. Formalising a novel and generic adversary model explicit to stealthy energy theft and benign anomalies in consumption and generation measurements.

2. Introducing a novel energy theft detection system defined by the synergy of an adaptive feature composition scheme and an smart meter classification component resulted by the aggregation of weather condition measurements and misconfiguration events over energy consumption and DRES deployments.

3. Constructing a self-learning process to enable our system to continuously and autonomously retrain based on instantly available measurements.

The rest of this chapter is structured as follows; Section 6.1 presents a generic model for mapping energy theft and misconfiguration events. Section 6.2 describes the methodology underpinning our detection system whereas Section 6.3 demonstrates our evaluation methodology. Section 6.4 evaluates the proposed soluton and demonstrates its ability to achieve high precision and accuracy in energy theft detection, whereas Section 6.5 concludes this work.

6.1 Smart grid & energy theft

6.1.1 System description

We consider an energy distribution network $G = \{A, N\}$ consisting of a set of consumers A distributed in several geographical regions and a set of low/medium voltage distribution buses N. Bidirectional data communication and power streams are used for energy transmission and distribution through corresponding power systems and networked data management components. Each consumer u_i in G is equipped with an smart meter to measure energy consumption. The consumption of a single u_i at a given hour $h \in H$ for a day $d \in D$ and month $m \in M$, is represented by $Ec_i(h, d, m)$. For this representation, $H = 1, 2, ..., 24$, $D = 1, 2, ..., 30$ and $M = 1, 2, ..., 12$ are defined as the set of hours within a day, the set of days in a month and the set of months in a year, respectively. We define a subset $B \subseteq A$ as a group of consumers owning and managing a DRES installation (e.g. domestic solar panels) as well as consuming power (i.e., prosumers). The energy produced by a single prosumer $i \in B$ in a given time period h, d, m is measured by a second smart meter and mapped as the function $Er_i(h, d, m) : \ Er_i(h, d, m) = 0 \ \ \forall i \notin B$.

In this context, energy theft activities result in energy losses defined as the difference between the energy supplied into a grid and the energy consumed under normal conditions (K. Zheng et al., 2018). Thus, the cumulative energy loss over a single time period h, d, m can be expressed as follows:

$$\begin{aligned} NTL(h, d, m) = \Delta Es(h, d, m) + \sum_{i=1}^{|B|} \Delta Er_i(h, d, m) \\ - \sum_{i=1}^{|A|} \Delta Ec_i(h, d, m) + \sum_{i=1}^{|N|} TL_i(h, d, m) \end{aligned} \tag{6.1}$$

where $Es(h, d, m)$ is the energy supplied by the utility provider to all individuals in A at a time interval h, d, m. Δ is the discrepancy in the smart meter measurements for the actual and reported readings of a single consumer/prosumer u_i due to the energy theft activities at time h, d, m, and TL is the transmission line losses caused by physical restrictions.

6.1.2 Energy theft and smart meter misconfiguration model

The primary assumption of this work is that prosumers and/or consumers can manipulate their consumption and/or generation measurements to report erroneous energy readings. Thus, in Eq. 6.1, we rule out discrepancies in the energy supplied

Table 6.1: Energy theft and smart meter misconfiguration functions where α, β, $\gamma(.)$, $\zeta(.)$, $\iota(.)$ and $\tau(.)$ are anomaly coefficients.

Type	Measurement	Function
Curtailment misconfiguration	Generation	$\Delta Er_i(h,d,m) = \alpha Er_i(h,d,m) \quad \forall\, Er_i(h,d,m) > 0$, where $\{\ \alpha \in \mathbb{R} \mid 0 \leq \alpha < 1\}$
Amplification misconfiguration	Consumption	$\Delta Ec_i(h,d,m) = \beta Ec_i(h,d,m) \quad \forall\, Ec_i(h,d,m) > 0$, where $\{\beta \in \mathbb{R} \mid \beta > 1\}$
Disconnect misconfiguration	Generation	$\Delta Er_i(h,d,m) = \text{NaN}$
	Consumption	$\Delta Ec_i(h,d,m) = \text{NaN}$
Total scaling theft	Generation	$\Delta Er_i(h,d,m) = \gamma(h,d,m) Er_i(h,d,m) \quad \forall\, Er_i(h,d,m) > 0$, where $\{\gamma(h,d,m) \in \mathbb{R} \mid \gamma > 1\}$
	Consumption	$\Delta Ec_i(h,d,m) = \zeta(h,d,m) Er_i(h,d,m) \quad \forall\, Ec_i(h,d,m) > 0$, where $\{\ \zeta(h,d,m) \in \mathbb{R} \mid 0 \leq \zeta(h,d,m) < 1\}$
Partial scaling theft	Generation	$\Delta Er_i(h,d,m) = \begin{cases} Er_i(h,d,m), & Er_i(h,d,m) \geq \iota \\ \iota, & Er_i(h,d,m) < \iota \end{cases}$ where $\{\iota > \in \mathbb{R} \mid \iota > \text{Min}\big(Er_i(1,d,m), Er_i(2,d,m), ..., Er_i(24,d,m)\big)\}$
	Consumption	$\Delta Ec_i(h,d,m) = \begin{cases} Ec_i(h,d,m), & Ec_i(h,d,m) \leq \tau \\ \tau, & Ec_i(h,d,m) > \tau \end{cases}$ where $\{\tau \in \mathbb{R} \mid \tau < \text{Max}\big(Ec_i(1,d,m), Ec_i(2,d,m), ..., Ec_i(24,d,m)\big)\}$
Off-peak theft	Generation	$\Delta E_r(h,d,m) = \begin{cases} \gamma(h,d,m) Er_i(h,d,m), & hs \leq h \leq he \mid Er_i(h,d,m) > 0 \\ Er_i(h,d,m), & otherwise \end{cases}$ where hs and he is the off-peak operating weather conditions for DRES.
On-peak theft	Consumption	$\Delta E_c(h,d,m) = \begin{cases} \zeta(h,d,m) Ec_i(h,d,m), & hb \leq h \leq hc \mid Ec_i(h,d,m) > 0 \\ Ec_i(h,d,m), & otherwise \end{cases}$ where hb and hc is the on-peak load hours.
Reply theft	Generation	$\Delta Er_i(h,d,m) = \text{Max}\big(Er_i(h-1,d,m), Er_i(h,d,m)\big)$
	Consumption	$\Delta Ec_i(h,d,m) = \text{Min}\big(Ec_i(h-1,d,m), Ec_i(h,d,m)\big)$
Stability theft	Generation	$\Delta Er_i(h,d,m) = \text{Max}\big(Er_i(1,d,m), Er_i(2,d,m), ..., Er_i(24,d,m)\big)$
	Consumption	$\Delta Er_i(h,d,m) = \text{Min}\big(Ec_i(1,d,m), Ec_i(2,d,m), ..., Ec_i(24,d,m)\big)$

by the utility provider having $\Delta Es(h,d,m) = 0$, since this measurement is assumed to be usually secure under a reliable communication link (Yip, Wong, et al., 2017a). However, the discrepancy in the smart meter generation and consumption readings in Eq. 6.1, Δ, may occur by both a theft-related activity or a non-malicious event, such as a misconfiguration. In general, energy losses defined in Eq. 6.1 relate to metering conditions that manifest as anomalous behaviour in the measurement of that particular meter. Such occurrences are common in instances of energy theft-related activities as well as smart meter misconfiguration incidents.

Therefore, to distinguish smart meter discrepancies, we present in Table 6.1 a taxonomy of smart meter anomaly function definitions, based on energy consumption and generation measurements. All functions mimic pragmatic characteristics of fraudulent or misconfigured smart meter patterns in terms of erroneously reported measurements. There anomaly functions are based on findings in the literature (Mahmoud et al., 2020; Yip, Wong, et al., 2017a; K. Zheng et al., 2018; M. Liu et al., 2020; David, 2021; Peppanen et al., 2015; Dai et al., 2021) and reflect a representative collection of common anomaly operations. According to Table 6.1,

misconfiguration and theft activities can partially or completely change the reported energy timeseries signal. Theft activities within consumption measurements target to decrease the monetary value of a consumer and thus they are mapped as a direct decrease in consumed energy. However, smart meter misconfigurations lead to unexpected increases in consumed energy. Theft functions exploiting DRES generation measurements with a goal for monetary gain feature an increase in the reported generated energy, while the misconfiguration of the DRES's smart meter is described as the curtailment of DRES energy back to the grid.

In more detail, the curtailment misconfiguration considers a scenario in which the misconfigured smart meter reports always less supplied energy than what the DRES deployment actually generated, for instance, 30% of the actual generated energy when $\alpha = 0.3$. However, in amplification misconfiguration, we assume that the misconfigured consumption smart meter consistently reports more than the actual amount of energy consumed by customers (i.e. $\beta > 1$). In the disconnect misconfiguration scenario, a smart metre loses its connection and is unable to continue transmitting energy measurements to the utility provider centre. This is one of the characteristics shared by both generation and consumption smart metres.

In case of the total scaling theft, we consider a scenario in which the total generation and consumption measurements are completely scaled by an attacker based on an arbitrary percentage, i.e. γ and ζ. For example, 140% of the actual generation measurements is reported by the attacker when $\gamma = 1.4$, while 50% of the actual consumed energy is reported when $\zeta = 0.5$. In the partial scaling theft scenario, only consumption measurements above a threshold ι and generation measurements under a threshold τ are scaled by the attacker. Therefore, a fraudulent prosumer sets a minimum reporting value for the DRES-based generation measurements sent to the main grid and a maximum value for the consumed energy every single day by a fraudulent consumer.

We also considered a case in which theft could not continuously occur, so there might be some discontinuous malicious reporting of the measurements during a certain period. For instance, in off-peak theft, the fraudulent prosumer reports 40% more power than that generated during off-peak hours, relating to the peak weather conditions in which the DRES operates. In on-peak theft, malicious consumers report 20% less than they consumed during on-peak load hours. However, in replay attacks, the fraudulent prosumer only reports the highest actual generation once it is reached; meanwhile, for consumption measurements, the attacker reports the minimum consumed energy. Finally, in the case of stability thefts, the attacker continuously reports the maximum generation for each day, and for the consumption measurement, the fraudulent consumer sends the minimum consumption of the day to take full benefit from the energy system business model.

6.2 Energy theft detection

As illustrated in Fig. 6.1, our system consists of two stages: (i) feature construction and, (ii) smart meter classification.

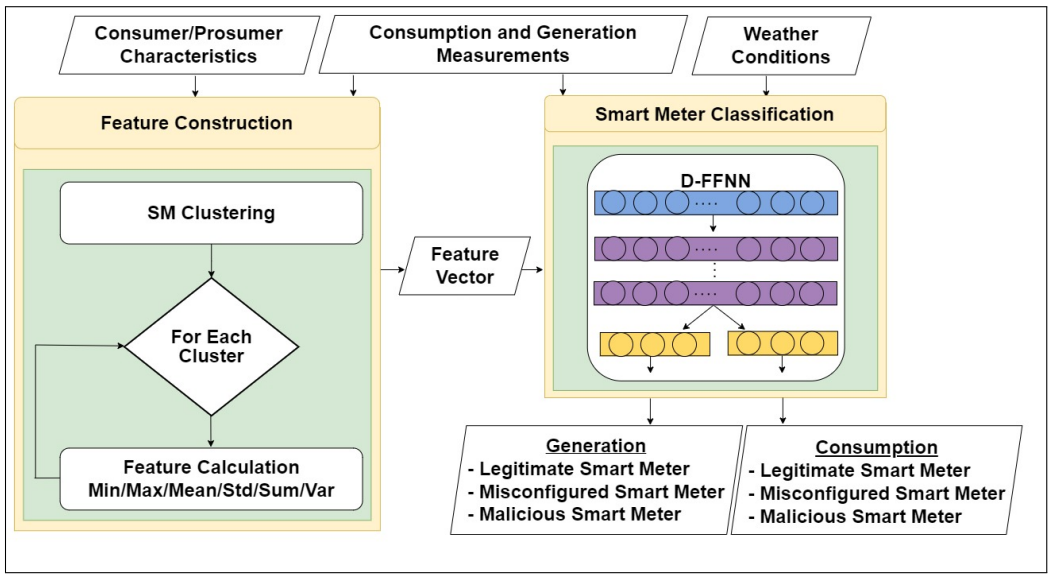

Figure 6.1: Data-flow of the proposed system.

6.2.1 Feature construction

The feature construction stage processes timeseries data from the infrastructure and builds extended feature sets. This is achieved by composing first order statistics (e.g., min/max, variance) of generation and consumption measurements for each group of consumers and prosumers. We consider correlated spatiotemporal behaviour across timeseries measurements by virtue of behavioral similarities in seasonal consumption and generation patterns.

For example, a peak in consumption pattern caused by air-conditioning demand during a hot wave could be observed across a large number of neighbouring consumers with similar characteristics. Similarly, prosumers managing solar panels can have a correlated generation pattern based on sunlight availability. It is thus feasible to establish a ground truth with respect to normal generation or consumption profiles. In particular, the feature construction module stage clusters smart meter using an incremental K-means algorithm, which partitions smart meters into k clusters based on a set of consumer/prosumer characteristics. These characteristics include geographical location, DRES physical characteristics, number of appliances, and tariff

agreement type. Clustering smart meters based on common characteristics allows the classification stage to extract common energy generation and consumption patterns emerging between consumers/prosumers within the same cluster (Angelos et al., 2011; D. Chen and Irwin, 2017).

Let an initial set of K-means $\left[\xi_1^{[1]}, \xi_2^{[1]}, ..., \xi_K^{[1]}\right]$, each consumer/prosumer $u_i \in G$ would group into a cluster whose mean is the shortest squared Euclidean distance as:

$$s_q^{[r]} = \left\{ u_i \ :|| \ u_i - \xi_q^{[r]} \ ||^2 \leq || \ u_i - \xi_j^{[r]} \ ||^2 \ \forall \ j \in [1, k] \right\} \tag{6.2}$$

In each iteration, the mean of the clusters can be updated as follows:

$$\xi_q^{[r+1]} = \frac{1}{\mid s_q^{[r]} \mid} \sum_{u_j \in s_q^{[r]}} u_j \tag{6.3}$$

Formally, the objective here is to minimise intra-cluster variance as:

$$\arg\min_{s} \sum_{q=1}^{k} \sum_{u \in s_q} || \ u - \xi_q \ ||^2 \tag{6.4}$$

where ξ_q is the mean of consumers and prosumers in s_q. The output of this process is a list of clusters $S = \left[s_1, s_2, ..., s_K\right]$ determining which cluster each individual $u_i \in G$ is grouped into and a list of the mean of each cluster $\Xi = \left[\xi_1, \xi_1, ..., \xi_K, \right]$ determining the mean of the individuals in each cluster. Once consumers and prosumers with correlated consumption and generation measurements are grouped, we calculate a set of variables representing regular consumption and generation patterns for the individuals within each cluster. Hence, for each cluster, we calculate the minimum (min), maximum (max), (var) variance, standard deviation (std), sum and mean of the generation and consumption measurements of a set of completely legitimate consumers/prosumers in that cluster. These variables provide different perspectives on the generated and consumed energy within that cluster, and overall they reflect the regular consumption and generation patterns for customers within that group. Thus, these features are preserved to serve as the ground truth of the regular generation and consumption patterns to support the detection process within each cluster.

6.2.2 Smart meter classification

Due to its ability to address multi-category problems and multiple tasks simultaneously, a dual deep feed forward neural network (D-FFNN) is defined in this module to determine whether each consumption and generation measurement is malicious,

misconfigured, or legitimate. The structure of the proposed D-FFNN comprises of an input layer with v neurons followed by l hidden layers, each with ne neurons, and finally, dual output layers, each with 3 neurons as per the category each sample is stratified (i.e., malicious, misconfigured, legitimate). The input layer sends the input data $X = \{x^{[1]}, x^{[2]}, ..., x^{[|X|]}\}$ to the hidden layers to extract features and understand patterns to facilitate producing a given category by each output layer.

Each $x^{[i]} \in X$ is an instance in the $\mathrm{v}-$dimensional feature space, i.e. $x^{[i]} = \big[x_1, x_2, ..., x_v\big]$. This feature space includes the reported consumed energy $Ec_i(h, d, m)$ and generated energy $Er_i(h, d, m)$ together with the features constructed in Section (6.2.1) and the weather conditions of i's geographical region for each $i \in G$ over the time slot h, d, m. The first output layer projects the category of the consumer's i consumption measurement $\hat{y}_c^{[i]}$ whereas the second output layer projects the category of the generation reading $\hat{y}_r^{[i]}$ at the time h, d, m. These two decisions indicate whether each smart meter of each consumer or prosumer is legitimate, malicious, or misconfigured. To train the D-FFNN, the input instances in X are mapped through the hidden layers from the input layer to the dual output layers as follows:

$$z^{[n]} = \sigma\big(\theta^{[n]} \cdot z^{[n-1]} + b^{[n]}\big) \quad \forall\, n \in [1, l+1] \tag{6.5}$$

where:

$$\begin{cases} z^{[0]} = x^{[i]} = \big[x_1, x_2, ..., x_v\big] \\ z^{[l+1]_1} = \hat{y}_c^{[i]} \\ z^{[l+1]_2} = \hat{y}_r^{[i]} \end{cases}$$

Here b represents a bias vector, θ is the connection weight, and $\sigma(\cdot)$ is a sigmoid function for the hidden layer and a softmax function for the dual output layers. The objective of the training process is to use a standard back-propagation to find b and θ. Algorithm 1 describes a workflow for the entire training process. In Algorithm 1, $\odot$ represents element-wise multiplication, ρ is a predefined learning rate, T is transpose operation, $\sigma'(\cdot)$ is the derivative of an activation function $\sigma(\cdot)$, and $\delta^{[n]}$ is the error in the layer n. The training process here is achieved by minimizing the dual objective function:

$$\arg\min_{\theta,b} J = \frac{1}{\mid X \mid} \sum_{i=1}^{|X|} \Big(L\big(\hat{y}_c^{[i]}, y_c^{[i]}\big) \quad + \quad L\big(\hat{y}_r^{[i]}, y_r^{[i]}\big) \Big) \tag{6.6}$$

where $y_c^{[i]}$ and $y_r^{[i]}$ represent the actual category corresponding to a sample $x^{[i]} \in X$ and $L(\cdot)$ is a three-class cross entropy function formulated as:

$$L(\hat{y}, y) = -\sum_{i=1}^{3} \eta_i \, y_i \log(\hat{y}_i) \tag{6.7}$$

where η represents an adjustment weight map for each category to force the detector to focus on the category where a larger learning loss occurs, resulting from an imbalance issue, to improve its performance. The initial D-FFNN defined by its trained parameters, i.e. θ abd b, is preserved to save the knowledge acquired during the learning process from the input data X. Thus, it can be used for detecting further measurements where each smart meter is listed in one of three groups – legitimate, malicious, or misconfigured – based on the results of the classification process.

Algorithm 1 D-FFNN training.

Initialise $\theta^{[n]}$ and $b^{[n]}$ randomly $\forall\, n \in [1, l+1]$
for each training sample $x^{[i]} \in X$ **do**
  for each layer $n \in [1, l+1]$ **do**
    Calculate $z^{[n]}(x^{[i]})$ using Equation 6.5
  end for
  Calculate $J(x^{[i]})$ using Equation 6.6 and Equation 6.7
  Calculate $\delta^{l+1}(x^{[i]}) = \nabla_z J(x^{[i]}) \odot \sigma'\big(\theta^{[l+1]} \cdot z^{[l]}(x^{[i]}) + b^{[l+1]}\big)$
  Calculate $\theta^{[l+1]} = \theta^{[l+1]} - \rho\, \delta^{[l+1]}(x^{[i]})\big(z^{[l]}(x^{[i]})\big)^T$
  Calculate $b^{[l+1]} = b^{[l+1]} - \rho\, \delta^{[l+1]}(x^{[i]})$
  for each hidden layer $n \in [l, 1]$ **do**
    Calculate

$$\delta^{[n]}(x^{[i]}) = \big((\theta^{[n+1]})^T \delta^{n+1}(x^{[i]})\big) \odot \sigma'\big(\theta^{[n]} \cdot z^{[n-1]} + b^{[n]}\big)$$

    Calculate $\theta^{[n]} = \theta^{[n]} - \rho\, \delta^{[n]}(x^{[i]})\big(z^{[n-1]}(x^{[i]})\big)^T$
    Calculate $b^{[n]} = b^{[n]} - \rho\, \delta^{[n]}(x^{[i]})$
  end for
end for

6.2.3 Self-learning operation

The self-learning operation of our detection system starts once a new batch of smart meter measurements is available. In this regard, new consumption and generation measurements are collected from the grid's consumers and prosumers, whose measurements may have been collected in the first data batch, or from new individuals who were connected recently to the power system. A generalised workflow of the self-learning operation is described in Algorithm 2. As illustrated in this algorithm, the system initially assigns each individual in the new batch to a corresponding cluster defined in the saved list of clusters S in Section 6.2.1.

However, if the new batch contains measurements from new consumers/prosumers, the squared Euclidean distance between these new individuals and the k-means in Ξ is measured. Subsequently, each new consumer/prosumer is assigned to the nearest cluster whose mean is the shortest distance, if this distance is smaller than a predefined threshold T_k. Otherwise, the system creates a new cluster for this new individual and updates the cluster set S and means set Ξ by adding the means of the recently created

cluster. We set the threshold T_k by referencing the longest distance between each individual and its cluster mean in the initial measurement batch. Once consumers and prosumers are clustered, the system calculates the set of features proposed in Section 6.2.1 from the newly available measurements to create the new input batch X' along with the weather data.

Algorithm 2 Self-learning operation.

Recall S and Ξ
Assign each consumer/prosumer u_i to its cluster s_q
for each new u_i **do**
    Find $\xi_q \in \Xi : \| u_i - \xi_q \|^2$ is the smallest
    if $\| u_i - \xi_q \|^2 < T_k$ **then**
        $s_q = s_q \cup u_i$
        Updated Ξ
    else
        Updated S
        Updated Ξ
    end if
end for
Construct features from each cluster
Collect weather condition measurements
Merge all measurements to create input data X'
Load D-FFNN
for each $x^{[i]} \in X'$ **do**
    $\hat{y}_r^{[i]}$ and $\hat{y}_c^{[i]} \leftarrow$ D-FFNN$(\Theta, x^{[i]})$
end for
Calculate AC_c and AC_r using Equation (6.8)
if $AC_c \leq T_c$ OR $AC_r \leq T_r$ **then**
Retrain D-FFNN with X' using Equation (6.5) to minimise the objective function in Equation (6.6)
end if
Save D-FFNN

Following the update of the new features based on newly available measurements, the previous version of the D-FFDD detection module is loaded such as to predict the consumption categories $\hat{y}_c$ and generation categories $\hat{y}_r$ in X'. The accuracy of this detection process is measured as follows:

$$AC = \frac{1}{3} \sum_{c=1}^{3} \frac{TP_c + TN_c}{TP_c + FN_c + FP_c + TN_c} \tag{6.8}$$

where TP are true positives, TN are true negatives, FN are false negatives, and FP are false positives.

As a result, we obtain two values AC_c and AC_r indicating the average number of correct predictions of $\hat{y}_c$ and $\hat{y}_r$, respectively, for all observations in X'. If one of the calculated values is less than the predetermined thresholds T_c and T_r, the batch is considered challenging, and the preserved D-FFDD is retrained automatically, using Equation (6.5), with the goal of minimizing the objective function in Equation (6.6). We set these thresholds by referencing arbitrary values around the accuracy of the training step of the system in the initial measurement batch. Similarly with the rest

of the parameters, the weights for the preserved D-FFNN will also be incrementally updated with the back-propagation as the new batch X' pass. This step is required such as the proposed classification module will adapt to observe the newly arrived generation and consumption measurements and self-optimise its own parameters.

6.3 Datasets and evaluation methodology

6.3.1 Datasets description

To validate our work, we utilise energy consumption and generation datasets collected in the power network of Australia's largest electricity provider, Ausgrid [1]. The dataset represents the generation and consumption measurements captured at a real installation of 300 different consumers and prosumers with rooftop solar panels from 1 July 2010 to 30 June 2013. However, in this work, we use only 139 individuals whose measurements were valid for the entire period. In addition to the consumption and generation smart meter measurements, the dataset includes information with respect to consumer/prosumer geolocation (e.g., postal codes) and solar panel capabilities (e.g., capacity).

As already mentioned, our system depends solely on weather conditions. For this purpose, we extracted available weather measurements from the World Weather Online API [2] and predictions of worldwide energy resources (POWER) project API [3] over the same observational period as that of the measurements obtained for Asugrid individuals.

6.3.2 Evaluation methodology

To demonstrate the effectiveness of our system, we conduct a performance comparison across four clustering algorithms named the density-based spatial clustering of applications with noise (DBSCAN), agglomerative nesting (AGNES), affinity propagation (AP), and fuzzy C-means clustering (FCM). We measure the **silhouette coefficient (SC)** score to evaluate whether individuals are clustered in well-defined groups. The SC is defined as:

$$SC = \frac{1}{\mid A \mid} \sum_{e=1}^{|A|} \frac{c(e) - o(e)}{\mathrm{Max}\big(c(e), o(e)\big)} \tag{6.9}$$

[1]Explore – Ausgrid Solar Home Electricity Data, Available:`https://www.ausgrid.com.au/Industry/Our-Research/Data-to-share/Solar-home-electricity-data`

[2]Explore – Weather API, Available:`https://www.worldweatheronline.com/developer/api/`

[3]Explore – Power Hourly API, Available:`https://power.larc.nasa.gov/api/pages/`

where $\mid A \mid$ is the total number of individuals in the grid, $c(.)$ is the average distance between a consumer/prosumer and other individuals in the same cluster, and $o(.)$ is the minimum average distance between that individual and all individuals belonging to other clusters.

Furthermore, we conduct a performance evaluation of various classification algorithms including both classic techniques (such as decision tree (DT), support vector (SVM) and K-nearest neighbours (K-NN)) and advanced methods (such as extreme gradient boosting (Xgboost)). This evaluation excludes advanced deep learning models such as long short-term memory networks (LSTM) because they require a two-dimensional feature vector, whereas our data is only one-dimensional. For this comparison, we utilise the following performance metrics:

1. **Precision (PR)** defined as:

$$PR = \frac{1}{3}\sum_{c=1}^{3}\frac{TP_c}{TP_c + TF_c} \tag{6.10}$$

2. **Recall (RE)** defined as:

$$RE = \frac{1}{3}\sum_{c=1}^{3}\frac{TP_c}{TP_c + FN_c} \tag{6.11}$$

3. **F1 Score (F1)** defined as:

$$F1 = 2 \times \frac{RE \times PR}{RE + PR} \tag{6.12}$$

4. **Area Under the Curve (AUC)** defined as:

$$AUC = \frac{1}{3}\sum_{i=1}^{3}\sum_{j>i}^{3}\frac{1}{2}\Big(\mathrm{BAUC}\big(i,j\big) + \mathrm{BAUC}\big(j,i\big)\Big) \tag{6.13}$$

where

$$BAUC\big(x,y\big) = \frac{\mathrm{Ranks} - \frac{AP}{2} \times (1 + AP)}{AP \times AN} \tag{6.14}$$

Here Ranks represents the sum of the ranks from class x, AP is the number of samples in class x and AN represents the number of samples in class y. The samples are arranged in ascending order based on the prediction of class i for ranking (Z. Zheng et al., 2018).

5. **Computational complexity:** to measure the inference time required to obtain classification decisions on test data.

It is worth mentioning that the computational complexity excludes the grid search process utilised to train and fine-tune hyper-parameters. It transforms a hyper-parameter domain into a grid and then traverses each point on the grid to obtain the optimal classifier parameters. Utilizing such a search strategy is straightforward, and the optimal search speed is quite reasonable. In addition, the optimal hyper-parameters are determined independently, enabling simultaneous optimization. Table 6.2 illustrates the results of the grid-search process for each classification algorithms.

Table 6.2: Optimal hyper-parameters of the classification algorithms.

Algorithm	Hyper-parameters
D-FFNN	l = 8, ne in hidden layer 1= 70, ne in hidden layer 2= 70, ne in hidden layer 3= 60, ne in hidden layer 4= 50, ne in hidden layer 5= 30, ne in hidden layer 6= 40, ne in hidden layer 7= 20, ne in hidden layer l= 4, Batch size = 32, Optimizer = adam, Learn rate = 0.001
DT	Maximum depth=12, Minimum samples split= 2, Minimum samples leaf= 2
SVM	Kernel= radial basis function, C= 1, Gamma= $0,2$
K-NN	Number of neighbors= 15
Xgboost	Number of estimators= 7, Maximum depth= 10

In addition, the **RE**, **PR**, **F1** and **AUC** are utilised to conduct a performance evaluation of the self-learning operation employed within the long-term theft detection process. The long-term theft detection process is described as the classifier trained on the initial batch training data is used to directly identify thefts and misconfigurations for the test set across other batches.

During our evaluation, we synthetically inject anomalous patterns within Ausgrid's dataset using the functions in Table 6.1 to emulate fraudulent and misconfigured samples. In order to avoid a data imbalance issue resulting from this procedure, a higher weight to the loss encountered by the samples associated with minor categories in Equation 6.7 is assigned. To note that we filter out instances of disconnect misconfgurations during the pre-processing stage. Evidently, such events demonstrate extremely large numbers of missing values in both generation and consumption measurements and they were affecting significantly the training phase. We also adjust the value of the solar panel smart meter to zero for a randomly chosen third of individuals to simulate simple consumers (i.e., not owning/managing a DRES).

Moreover, we group the Ausgrid dataset by year to simulate a scenario in which smart meter measurements were presented continuously over time. Nevertheless, to simulate a scenario in which new individuals join the grid, we removed the measurements of ten arbitrarily chosen individuals from the first batch, and reintroduce them

incrementally across batches. Each batch is split into training and testing sets, with a ratio of 70 : 30 respectively. In order to avoid bias issues here, we employ a 10-fold cross-validation scheme. This scheme divides each patch's sample into 10 distinct folds of approximately equivalent size (without repetition). One fold is used as a test set, while the other folds are utilised as a training set. This process is repeated 10 times, and the average performance across all 10 repetitions of the testing set is then calculated for consideration. We then normalise the training and test data incrementally to transform all values of the features into a single scale with unit variance and mean of zero. However, categorical time series are encodeed using a binary encoder.

6.4 Results

Following the evaluation methodology presented earlier, the produced outputs in Fig. 6.2 indicate that the k-means formulation achieved the highest SC score (i.e., SC=0.44). Thus, we utilise its capabilities for the proposed detection system. In addition, the D-FFNN, formulation performed better than all classification algorithms in detecting malicious and misconfigured meters as demonstrated in Table 6.3. With respect to generation measurements, the D-FFNN scheme recorded a precision of 0.92, 0.90, 0.85, 0.84 and 0.91 higher than that of DT, SVM, K-NN and Xgboost respectively. This was similar to the consumption measurements, where the D-FFNN outperformed the other classifiers, as it recorded a 72% precision accuracy, while the DT, SVM, K-NN and Xgboost achieved 0.68, 0.71, 0.59 and 0.68 respectively.

Table 6.3: Detection performance of the smart meter classification module using different algorithms.

Algorithm	Performance Parameter							
	Consumption				Generation			
	PR	RE	F1	AUC	PR	RE	F1	AUC
D-FFNN	0.72	0.69	0.70	0.73	0.92	0.92	0.92	0.91
DT	0.68	0.61	0.64	0.68	0.90	0.89	0.89	0.88
SVM	0.71	0.69	0.69	0.69	0.85	0.85	0.85	0.84
K-NN	0.59	0.49	0.53	0.60	0.84	0.84	0.84	0.83
Xgboost	0.68	0.65	0.66	0.67	0.91	0.90	0.90	0.89

Evidently, the D-FFNN superiority over DT, SVM, K-NN and Xgboost in generation and consumption measurements were uniform, even when RE, F1 and AUC scores were measured. We argue, that the D-FFNN formulation is superior due to its ability to capture hidden patterns in the weather condition data as well as the constructed features in Section 6.2.1. Higher detection performance in terms of RE, PR, F1 and AUC was observed particularly for the generation measurements as depicted in Table 6.3. The higher performance is attributed to the variables distilled

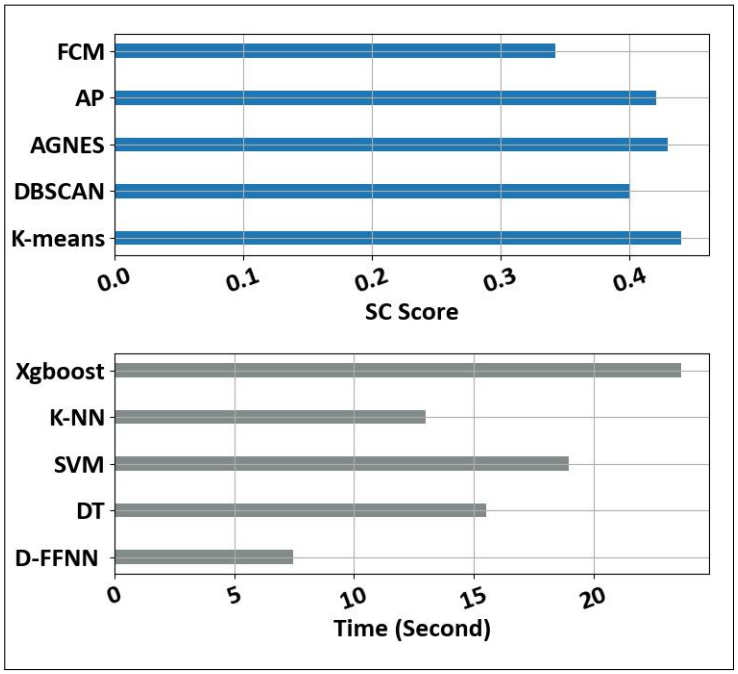

Figure 6.2: SC score of different clustering algorithms and comparison of the computational complexity time.

by the original solar panel capacity feature that was instrumental in profiling prosumer normal behaviour with respect to generation. Nonetheless, Ausgrid's dataset lacks consumer characteristics (e.g., number of rooms, appliances) that can contribute better to profiling consumers with similar consumption patterns. Therefore, the use of additional features in the smart meter clustering process to identify customers with similar patterns is necessary to improve the performance of the classification process in consumption measurements.

Apart from high precision accuracy, the D-FFNN formulation also operates with relatively lower computational time compared to other schemes as depicted in Fig. 6.2[4]. Arguably, this outcome revolves around the fact that the rest of the formulations required independently trained models explicit to either generation or consumption measurements incurring extensive computational overheads. This demonstrates the efficacy of using a dual, deep learning technique instead of conventional techniques in our detection system, as we need to train one model with two outputs to address both tasks simultaneously.

Deteriorated performance over time was observed in the long-term detection

[4]On 64-bit Windows operating system with an Intel Core i7 (7*th* Gen) CPU with a 2.80 GHz clock cycle and 32 GB of RAM.

process, where the 2010 training data is used to detect thefts and misconfigurations across 2011, 2012 and 2013 as depicted in Fig. 6.3. The impact on accuracy performance is a result of the change in data distribution properties across the batches. Consequently, such change misleads the detection system over the years and results in further detection errors.

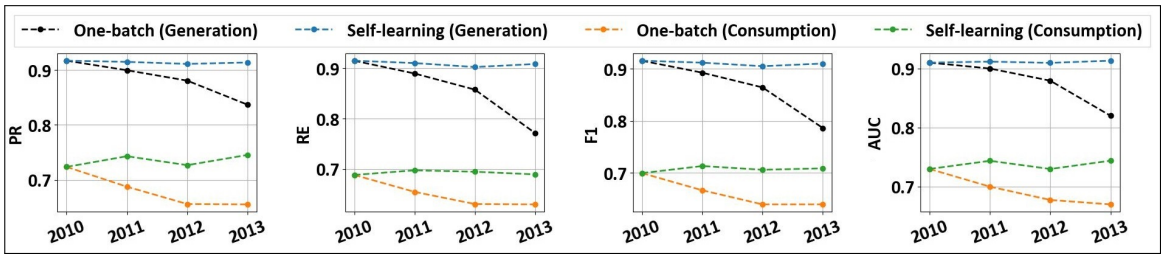

Figure 6.3: Long-term detection performance.

The main cause for such a change is attributed via including entirely new individuals whose measurement patterns vary from the patterns included in the first batch's training data. Therefore, the detection system fails to identify the measurements of consumers/prosumers who have recently connected to the network and decide whether they are legitimate, malicious, or misconfigured. Even in cases where no new individuals are linked with the grid, the generation and consumption measurements of the same individuals usually have non-stationary properties, so the distribution of the data also varies across batches. The non-stationary properties in the consumption measurements are caused by changes in consumption habits, for example, installing eco-friendly equipment that reduces energy consumption (Fekri et al., 2021), while the non-stationary nature of the generation measurements is usually caused by changes in the weather conditions over many years (Staffell and Pfenninger, 2018).

However, as depicted in Table. 6.4, whilst the detection accuracy of the consumption and generation measurements are less than the predefined thresholds in the 2011 data batch, the system considers the batch challenging and self-optimises by retraining on that batch. As a result, the performance of the proposed system is improving based on the test data of the 2011 batch (see Fig. 6.3). The same process was also carried out for the years 2012 and 2013 to ensure validation.

6.5 Summary

Modern energy theft techniques exploit the highly distributed nature of the modern smart grid and cause significant financial loss to energy providers. Hence, tracking such events is critical but also challenging due to the diversity of the composite

Table 6.4: Accuracy of the smart meter classification module in the long-term detection process.

Batch	Consumption	Generation	Consumption Threshold	Generation Threshold
2010	0.72	0.92	0.71	0.91
2011	0.67	0.90	0.71	0.91
2012	0.67	0.87	0.71	0.91
2013	0.70	0.90	0.71	0.91

attack vectors triggering them where faults promote the same properties as theft. In this paper, we propose a self-learning system that can distinguish energy theft from faults with the joint use of consumption and generation measurements as well as openly available weather information. The outcomes of an extensive and comparative evaluation over real measurements reveal that the introduced scheme can reach over 90% of accuracy under the D-FFNN formulation and with relatively low computational overheads. Its joint use with other ML techniques under the proposed methodology that can provide for the analysis of online measurement streams can adequately adapt over varying properties of theft or misconfiguration scenarios. It can thus benefit the design of next-generation energy theft detection systems.

Chapter 7

Conclusions and Future Directions

This chapter provides a summary of this thesis conducted on designing a practical and data-driven framework considering the advanced characteristics of modern energy grids. In addition, it highlights a number of unresolved issues that could be the subject of future research in this field.

7.1 Conclusion

Smart grids are the result of the convergence of legacy energy system components and advanced information and communication technologies. Smart grids are viewed as innovative, next-generation power systems that can ensure cost-effective decarbonisation of the energy sector. They are accomplished by collecting data and measurements from various elements of grid infrastructures (i.e. generation, T&D, and end-user) in order to extend grid monitoring, observation and control. However, the development of such systems brings it is accompanied by a wide range of security threats that increase the opportunities for energy theft-related attacks. In addition to the financial gain that it brings for a malicious actor, energy theft can have a direct effect on the overall resilience and safety of smart energy ecosystems. Associated with the invention of the smart grid, the primary focus of this thesis is to design a practical and data-driven framework for energy theft detection in the modern energy grid environment.

To achieve this, this thesis addresses the following research questions: 1) how does the introduction of smart grids enable larger attack vectors that provide a basis for energy theft?; and 2) how do we design a data-driven framework for detecting energy theft? The latter question can be divided into the following sub-questions: a) how can we leverage diverse sources of measurement to identify energy theft attacks in DRES?; and b) how can we devise a generic method to accurately detect energy theft in scalable smart grids?

With regard to the first question, Chapter 2 conducts a comprehensive review of

energy theft attacks and detection methods for smart grid systems. In this regard, the chapter begins with a discussion of smart grid components in the energy supply chain, with an emphasis on data communication, as well as the pillars for assessing grid effectiveness. The impact of energy theft on the smart grid is then evaluated by analysing how demand, supply and generation data manipulation can facilitate energy theft activities targeting energy grids. Examples of data-driven and data-agnostic energy theft attacks and their enabling techniques are then discussed. Additionally, this chapter categorises the main research studies, addressing the aspect of energy theft detection, and summarises the experimental approaches applied in this research. This chapter concludes by highlighting a number of open issues and challenges in the field of energy theft detection in modern energy ecosystems.

In light of this chapter, the answer to the first research question is that the smart grid paradigm, characterized by the integration of the data collection infrastructure (cyber infrastructure) running through smart grid components (including generation, T&D and end-user infrastructures), enables additional energy theft activities in modern energy systems, while also suggesting promising detection solutions. Therefore, the smart grid paradigm represents a double-edged sword for the security and forensic domains of modern energy systems.

On the one hand, the integrated data collection infrastructure within modern energy grids enables a wide range of applications, such as energy trading platforms. These applications rely on inherently vulnerable networked environments, such as advanced metering systems. Malicious actors exploit these vulnerabilities by manipulating communication and energy measurement-related data generated and stored by networked metering, management and control devices. By tampering with data integrity and energy measurement precision, these actors can report false information, leading to financial gain through energy theft or fraudulent trading behaviours. Additionally, this can disrupt the grid's reliability, potentially causing power disruptions or supply-demand imbalances.

On the other hand, data-driven strategies have enormous potential for detecting energy theft-related activities across smart grid infrastructures. Utilizing measurements and information collected by an integrated cyber-infrastructure infrastructure, utility providers are able to develop data-driven energy theft detection strategies. Such strategies comprise an algorithmic solution that emphasizes data deviations related to aspects such as metering and billing. Hence, these detection schemes put significant emphasis on analysing data patterns with a variety of statistical tools, and the vast majority employ machine learning techniques.

Regarding the second research question, Chapter 3 proposes a widely applicable theoretical framework for data-driven energy theft detection. According to the study conducted in Chapter 2, we have noticed that there are still some uncertainties over the architecture, components and resources of data-driven energy theft detection

strategies. Motivated by this observation, our second contribution in Chapter 3 proposes a general theoretical framework for a data-based energy theft detection process. It is based on in-depth comprehension of the energy theft problem and the detection requirements for modern energy grids. The proposed framework is a complementary four-dimensional structure. These four dimensions are infrastructure, measurement monitoring control (MMC), operation and end-users. These are evolving domains that allow utility operators to continuously identify theft activities and thus optimise their detection procedures.

By utilizing the theoretical framework proposed in this chapter, we can develop two distinct data-driven detection approaches for identifying theft activities in various scenarios. These data-driven detection contributions, derived from the proposed theoretical framework are: i) a predictive energy theft detection approach for distributed energy sources (DRES); and ii) an adaptive energy theft detection approach for consumption and generation smart meters.

The first proposed approach, a predictive energy theft detection method for DRES theft scenarios, is discussed in both Chapter 4 and Chapter 5. In Chapter 4, a SCADA-agnostic energy modelling system for DRES is proposed. This DRES profiling method enables automated feature selection and the tuning of regression models based on machine learning, facilitating adaptation to various measurement inputs. Building on the work in Chapter 4, Chapter 5 introduces a predictive SCADA-agnostic energy theft detection approach for DRES-based scenarios. The proposed detection approach involves two algorithms: i) a SCADA-agnostic DRES profiling scheme operating purely on third-party and widely available weather measurements; and ii) a classification scheme relying on DRES profiling that is able to classify theft detection events. In addition, this chapter provides a formalised approach for describing DRES-based adversaries with the objective of energy theft.

According to the performance evaluations of the data-driven approaches proposed in these two chapters, the answer to research question $2a$ is that using freely available third-party weather data can be effectively applied to detect energy theft-based DRES scenarios. Hence, in Chapter 4, the SCADA-agnostic energy profiling performance of regression-based models employing SCADA measurements can be adequately matched on a large scale using freely available third-party weather data. Moreover, the SCADA-agnostic approach proposed in Chapter 4 is achieved with a minimal set of weather parameters and at a lower computational cost than SCADA-based energy profiling. This opens the door to independent and cost-effective power generation profiling to support various planned smart grid applications, including the detection of malicious actors. In addition, using evaluations based on an energy profile model and third-party weather data, Chapter 5 demonstrates that the proposed predictive detection strategy can detect fraudulent DRES measurements with relatively low computing costs and a high average accuracy rate. These evaluations are conducted

by employing freely available third-party weather measurements from actual solar and wind energy deployments in Australia and France, respectively. These outcomes show that the predictive approach presented in this chapter is a viable and cost-effective solution for data-driven energy theft detection in the DRES context.

The second proposed approach derived from our theoretical framework in Chapter 3, which is a an adaptive energy theft detection approach for consumption and generation smart meters, is discussed in Chapter 6. The proposed energy theft detection approach in this chapter is a self-learning system that can differentiate between energy theft and misconfigurations by combining consumption and generation measurements with openly available weather data. The proposed detection strategy is defined by the synergy of: i) an adaptive feature composition scheme and ii) a smart meter classification component working on the aggregation of weather condition measurements and misconfiguration events over DRES and consumption deployments. In addition, this chapter introduces the formalisation of a novel and generic adversary model explicit to stealthy energy theft causing benign anomalies in consumption and generation measurements.

Based on performance evaluations of the energy theft detection approach proposed in this chapter, the answer to research question 2*b* is that by employing the synergy of adaptive data-driven methods, we can develop a technique to accurately detect energy thefts in scalable smart grids. The approach presented in this chapter can achieve relatively high accuracy with minimal computing overheads for detecting energy theft activities and misconfiguration instances across generation and consumption smart meters. Furthermore, the proposed methodology is capable of continuous and autonomous retraining, utilizing instantly available measurements. Therefore, it can contribute to the development of next-generation energy theft detection systems in scalable smart energy grids.

7.2 Future directions

Future research on data-driven energy theft detection in smart grid environments will focus on enabling the next generation of energy theft detection, in response to the emergence of advanced energy trading market applications (such as ancillary services and virtual power plants). There are two directions in which the work presented in this thesis can be extended for the next generation of energy theft detection: i) developing a formalised approach to describing how these applications facilitate energy theft attacks, and ii) devising a data-driven strategy for detecting these energy theft attacks.

For the former direction, our energy theft categorisation model introduced in Chapter 2 can be applied. We argue that our energy theft model can be used to derive

a set of discrete functions that describe the core concept behind any given theft attack. The inter-dependent variables of these proposed functions can be adapted according to the infrastructures containing the targeted applications; this will provide a formal definition of how the manipulation of these applications can increase the non-technical energy loss formulated by our energy theft model. Hence, these functions rely on the variable-specific manipulations of a malicious actor, based on the intrinsic properties of the system and/or the network components of the targeted applications.

In this direction, Chapters 5 and 6 detail specific applications of the proposed energy theft model. On the basis of our energy theft model, Chapter 5 presents a formalised method for describing DRES-based adversaries with the goal of energy theft. The proposed DRES-based theft model illustrates how effectively DRES owners can exploit the existing business model for financial gain, thereby increasing non-technical energy loss for a given TSO. In addition, in Chapter 6, we formalise a generic adversary model based on our energy theft model that is explicitly for stealthy energy theft, which causes benign anomalies in consumption and generation measurements. These proposed formulas describe how malicious actors exploit smart metering system vulnerabilities to increase non-technical energy losses for financial gain.

For the latter of the above directions, our proposed theoretical framework described in Chapter 3 can be applied to instantiate data-driven approaches to energy theft detection that are explicitly applicable to the next generation of energy theft scenarios. The following paths can be considered in this regard:

1. Including additional features within the detector construction sub-process of the operation dimension of the proposed theoretical framework in order to improve the overall detection accuracy of the detection approaches. These features could be, for instance: (i) information regarding grid consumers, such as the number of appliances and rooms; (ii) information regarding the DRES, such as accurate longitude and latitude; and (iii) information on underlying communication architectures. The integration of additional features could provide greater insights into generation and consumption patterns throughout the energy flow, thereby enhancing the ability of data-driven detectors to detect anomalies related to energy theft activities. However, a trade-off should be made between the efficiency benefit and the issue of over-fitting. Hence, the addition of features can lead to a detection strategy that is specifically tailored to suit particular data conditions and settings, limiting its generalisability.

2. Conducting simulated testbed-based evaluations within the operation dimension of the proposed theoretical framework in order to gain insights into the performance of detection approaches. While evaluating energy theft detectors using real-world datasets provides a more accurate understanding of smart grid environments, such datasets are either not commonly available or primarily

created for specific projects. We argue that simulation analyses can provide significant exposure to settings and scenarios (e.g., various DRES billing and trading approaches) that are unavailable within the current real-world validations. In this regard, MATLAB Simulink (MathWorks, 2021) can be used to create virtual energy grids. This is a graphical programming environment that permits the performance of individual grid components to be modelled and simulated in order to generate grid-wide measurements. These measurements can be used in the development of data-driven detection algorithms.

3. Incorporating an online-learning theft detection sub-process within the operation dimension of the proposed theoretical framework in order to detect unseen (unknown) theft attacks. Within smart grid deployments, malicious actors frequently introduce innovative techniques to manipulate the smart grid business model and gain financial benefits. These new theft behaviours are referred to as zero-day theft attacks, and they may not be anticipated during the training phase of theft detectors, as they have never occurred before (Shaaban et al., 2021). Therefore, it is essential to develop online detection algorithms that can tolerate zero-day energy theft attacks and gradually identify new theft patterns across the smart grid. In such a situation, a theft attack repository can be utilised to collect unknown malicious behaviours. These unknown attacks are then combined with a portion of previous theft behaviour seen during detector training, and fed into the classification technique for online learning of these new theft events.

4. Leveraging quantum machine learning strategies within the proposed theoretical framework's operation dimension's detector construction sub-process. In contrast to traditional computers based on the physical implementation of the 0 and 1 states, quantum computers employ the combination of two quantum states $|0\rangle$ and $|1\rangle$ in a qubit to perform several computational processes concurrently (Schuld, Sinayskiy, and Petruccione, 2015). Recent studies have investigated the possibility of employing quantum computing to improve machine learning techniques, therefore the use of such quantum machine learning algorithms to develop advanced data-driven energy theft solutions is a promising notion. The benefits anticipated from these strategies include an improvement in theft detector learning efficiency, meaning that the same detection performance can be achieved with fewer training data or more simple architectures. In addition, the use of quantum machine learning techniques can enhance the computational overheads by obtaining faster theft identification across scalable energy systems.

5. Incorporating a privacy-preserving technique within the detector construction sub-process of the operation dimension of the proposed theoretical framework

in order to protect the private measurements utilised in data-driven detection processes. Such private information could be purchased by marketing companies seeking to target individuals who are likely to be interested in their products. In addition, criminals gaining access to this information can use knowledge about a resident's energy use and generation practices to better organise their attacks (S. Salinas, Ming Li, and P. Li, 2012). In addition to approaches for preserving the identities of grid users (such as pseudonyms and anonymization algorithms) and for protecting users' data (such as data obfuscation algorithms and homomorphic encryption), a federated learning-based detection model can be used to maintain grid users' privacy. In federated learning, theft detectors are trained on a decentralised edge node, which retains all the training data (Yan and H. Wen, 2021). Generally, privacy-preserving strategies are costly since they require substantial computational resources (Ahmed et al., 2022). To achieve a low total cost, energy theft detection systems should integrate privacy protection strategies and account for their associated costs.

6. Developing a top-down energy theft detection system within the operation dimension in order to encompass the entire life cycle of energy data through the MMC dimension of the proposed theoretical framework. In the field of data-driven energy detection, system integration demand is always one of the primary considerations (Jindal, Dua, et al., 2016; Yan and H. Wen, 2021). Hence, it would be ineffective for providers at different levels of aggregation to deploy and synchronise monolithic energy theft detection techniques due to their demanding processing needs. To address this issue, multiple input multiple output neural networks can be utilised to construct energy theft detectors. Such networks receive multiple measurements as inputs from each of the three infrastructures of the smart grids and return three outputs simultaneously, where each output is an identification of energy theft in the corresponding infrastructure. In addition to decreasing the computing demands of deploying detection procedures, this strategy delivers benefits such as improved data efficiency, reduced over-fitting through the use of shared representations, and rapid learning by using auxiliary information (Crawshaw, 2020).

References

Abedinia, Oveis et al. (2020). "Improved EMD-based complex prediction model for wind power forecasting". In: *IEEE Transactions on Sustainable Energy*.

Afrin, Sabrina and Sumita Mishra (2016). "An anonymized authentication framework for smart metering data privacy". In: *2016 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. IEEE, pp. 1–5.

Ahmed, Mohsin et al. (2022). "Energy theft detection in smart grids: taxonomy, comparative analysis, challenges, and future research directions". In: *IEEE/CAA Journal of Automatica Sinica* 99, pp. 1–23.

Aldegheishem, Abdulaziz et al. (2021). "Towards sustainable energy efficiency with intelligent electricity theft detection in smart grids emphasising enhanced neural networks". In: *IEEE Access* 9, pp. 25036–25061.

Alkhresheh, Ashraf, Mutaz AB Al-Tarawneh, and Mohammad Alnawayseh (2022). "Evaluation of Online Machine Learning Algorithms for Electricity Theft Detection in Smart Grids". In: *Evaluation* 13.10.

Alladi, Tejasvi et al. (2019). "Blockchain in smart grids: A review on different use cases". In: *Sensors* 19.22, p. 4862.

Althobaiti, Ahlam, Anish Jindal, and Angelos K Marnerides (2020). "Scada-agnostic power modelling for distributed renewable energy sources". In: *2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)*. IEEE, pp. 379–384.

Althobaiti, Ahlam, Anish Jindal, Angelos K Marnerides, and Utz Roedig (2021). "Energy Theft in Smart Grids: A Survey on Data-Driven Attack Strategies and Detection Methods". In: *IEEE Access* 9, pp. 159291–159312.

Althobaiti, Ahlam, Anish Jindal, and Angelos K. Marnerides (2021). "Data-Driven Energy Theft Detection in Modern Power Grids". In: *Proceedings of the Twelfth ACM International Conference on Future Energy Systems*. e-Energy '21. New York, NY, USA: Association for Computing Machinery, pp. 39–48.

Althobaiti, Ahlam, Charalampos Rotsos, and Angelos K Marnerides (2023). "Adaptive Energy Theft Detection in Smart Grids Using Self-Learning With Dual Neural Network". In: *IEEE Transactions on Industrial Informatics*.

Amin, Massoud and John Stringer (2008). "The electric power grid: Today and tomorrow". In: *MRS bulletin* 33.4, pp. 399–407.

Angelos, Eduardo Werley S et al. (2011). "Detection and identification of abnormalities in customer consumptions in power distribution systems". In: *IEEE Transactions on Power Delivery* 26.4, pp. 2436–2442.

Arcos-Aviles, Diego et al. (2016). "Fuzzy logic-based energy management system design for residential grid-connected microgrids". In: *IEEE Transactions on Smart Grid* 9.2, pp. 530–543.

Ashok, Aditya, Manimaran Govindarasu, and Venkataramana Ajjarapu (2018). "Online detection of stealthy false data injection attacks in power system state estimation". In: *IEEE Transactions on Smart Grid* 9.3, pp. 1636–1646.

Ashraf, Muhammad Mansoor et al. (2022). "FedDP: A Privacy-Protecting Theft Detection Scheme in Smart Grids Using Federated Learning". In: *Energies* 15.17, p. 6241.

Ashrafuzzaman, Mohammad, Yacine Chakhchoukh, et al. (2018). "Detecting stealthy false data injection attacks in power grids using deep learning". In: *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)*. IEEE, pp. 219–225.

Ashrafuzzaman, Mohammad, Saikat Das, et al. (2020). "Detecting stealthy false data injection attacks in the smart grid using ensemble-based machine learning". In: *Computers & Security* 97, p. 101994.

Aydin, Zafer and V Cagri Gungor (2018). "A Novel Feature Design and Stacking Approach for Non-Technical Electricity Loss Detection". In: *2018 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)*. IEEE, pp. 867–872.

Banshwar, Anuj et al. (2017). "Renewable energy sources as a new participant in ancillary service markets". In: *Energy strategy reviews* 18, pp. 106–120.

Basumallik, Sagnik et al. (2017). "Impact of false data injection attacks on PMU-based state estimation". In: *2017 North American Power Symposium (NAPS)*. IEEE, pp. 1–6.

BBC News (2019). *Ransomware hits Johannesburg electricity supply*. URL: `https://www.bbc.co.uk/news/technology-49125853`.

Bhattarai, Bishnu P et al. (2019). "Big data analytics in smart grids: state-of-the-art, challenges, opportunities, and future directions". In: *IET Smart Grid* 2.2, pp. 141–154.

Bihl, Trevor J and Salam Hajjar (2017). "Electricity theft concerns within advanced energy technologies". In: *2017 IEEE National Aerospace and Electronics Conference (NAECON)*. IEEE, pp. 271–278.

Bor, Martin C. et al. (2019). "Adversarial Machine Learning in Smart Energy Systems". In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy '19. Phoenix, AZ, USA: Association for Computing

Machinery, pp. 413–415. ISBN: 9781450366717. DOI: 10.1145/3307772.3330171. URL: https://doi.org/10.1145/3307772.3330171.

Burgess, Patrick R and Funlade T Sunmola (2021). "Prioritising requirements of informational short food supply chain platforms using a fuzzy approach". In: *Procedia Computer Science* 180, pp. 852–861.

Burke, Matthew J. and Jennie C. Stephens (2018). "Political power and renewable energy futures: A critical review". In: *Energy Research & Social Science* 35. Energy and the Future, pp. 78–93. ISSN: 2214-6296. DOI: https://doi.org/10.1016/j.erss.2017.10.018.

Buzau, Madalina-Mihaela et al. (2018). "Detection of non-technical losses using smart meter data and supervised learning". In: *IEEE Transactions on Smart Grid*.

Carr, Darragh and Murray Thomson (2022). "Non-technical electricity losses". In: *Energies* 15.6, p. 2218.

CEN-CENELEC-ETSI, Smart Grid Coordination (2012). "Group.(2012)". In: *Smart Grid Reference Architecture*, pp. 1–107.

Cetinkaya, Oktay and Ozgur B Akan (2017a). "Electric-field energy harvesting from lighting elements for battery-less internet of things". In: *IEEE Access* 5, pp. 7423–7434.

— (2017b). "Electric-field energy harvesting in wireless networks". In: *IEEE Wireless Communications* 24.2, pp. 34–41.

Chang, Keun-Su et al. (2012). "Electric field energy harvesting powered wireless sensors for smart grid". In: *Journal of Electrical Engineering and Technology* 7.1, pp. 75–80.

Chen, Dong and David Irwin (2017). "Sundance: Black-box behind-the-meter solar disaggregation". In: *Proceedings of the eighth international conference on future energy systems*, pp. 45–55.

Chen, Tianqi and Carlos Guestrin (2016). "Xgboost: A scalable tree boosting system". In: *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Chen, Yize, Yushi Tan, and Deepjyoti Deka (2018). "Is machine learning in power systems vulnerable?" In: *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. IEEE, pp. 1–6.

Cheng, Gang et al. (2022). "A highly discriminative detector against false data injection attacks in AC state estimation". In: *IEEE Transactions on Smart Grid* 13.3, pp. 2318–2330.

Cheng, Xu et al. (2022). "Wind turbine blade icing detection: A federated learning approach". In: *Energy* 254, p. 124441.

Chin, Wen-Long, Ya-Hsuan Lin, and Hsiao-Hwa Chen (2016). "A framework of machine-to-machine authentication in smart grid: a two-layer approach". In: *IEEE Communications Magazine* 54.12, pp. 102–107.

Cintuglu, Mehmet Hazar et al. (2016). "A survey on smart grid cyber-physical system testbeds". In: *IEEE Communications Surveys & Tutorials* 19.1, pp. 446–464.

Cody, Christa, Vitaly Ford, and Ambareen Siraj (2015). "Decision tree learning for fraud detection in consumer energy consumption". In: *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. IEEE, pp. 1175–1179.

Collier, Steven E (2017). "The emerging enernet: Convergence of the smart grid with the internet of things". In: *IEEE Industry Applications Magazine* 23.2, pp. 12–16.

Crawshaw, Michael (2020). "Multi-task learning with deep neural networks: A survey". In: *arXiv preprint arXiv:2009.09796*.

Czechowski, Robert and Anna Magdalena Kosek (2016). "The most frequent energy theft techniques and hazards in present power energy consumption". In: *2016 Joint Workshop on Cyber-Physical Security and Resilience in Smart Grids (CPSR-SG)*. IEEE, pp. 1–7.

Dai, Yanjie et al. (2021). "Smart Electricity Meter Reliability Analysis Based on In-service Data". In: *2021 4th International Conference on Energy, Electrical and Power Engineering (CEEPE)*. IEEE, pp. 143–147.

dansie, M. (2013). *Free Electricity From Thin Air*. Accessed on: March 2020. URL: `http://revolutiongreen.%20com/free-electricity-from-thin-air/`.

David, Leonardo (2021). *Solar Panels Underperforming? Here's How to Fix Common Issues*. Accessed: 2022-01-22. URL: `https://www.ecowatch.com/solving-solar-panel-output-issues-2655223014.html`.

Diamantoulakis, Panagiotis D, Vasileios M Kapinas, and George K Karagiannidis (2015). "Big data analytics for dynamic energy management in smart grids". In: *Big Data Research* 2.3, pp. 94–101.

Din, Ghulam Mohi Ud, Andreas U Mauthe, and Angelos K Marnerides (2018). "Appliance-level short-term load forecasting using deep neural networks". In: *2018 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, pp. 53–57.

Dineshkumar, K, Prabhu Ramanathan, and Sudha Ramasamy (2015). "Development of ARM processor based electricity theft control system using GSM network". In: *2015 international conference on circuits, power and computing technologies [ICCPCT-2015]*. IEEE, pp. 1–6.

EIA, Electricity - (2016). *International Energy Outlook 2016*. URL: `https://www.eia.gov/outlooks/ieo/pdf/electricity.pdf.%20Accessed%20on:%20March%202019`.

Ekanayake, Janaka B et al. (2012). *Smart grid: technology and applications*. John Wiley & Sons.

El Mrabet, Zakaria et al. (2018). "Cyber-security in smart grid: Survey and challenges". In: *Computers & Electrical Engineering* 67, pp. 469–482.

Engebretson, Patrick (2013). *The basics of hacking and penetration testing: ethical hacking and penetration testing made easy*. Elsevier.

EPA (2018). *Centralized Generation of Electricity and its Impacts on the Environment*. Accessed on: March 2020. URL: https://www.epa.gov/energy/centralized-generation-electricity-and-its-impacts-environment.

Esmalifalak, Mohammad et al. (2017). "Detecting stealthy false data injection using machine learning in smart grid". In: *IEEE Systems Journal* 11.3, pp. 1644–1652.

Farhangi, Hassan (2009). "The path of the smart grid". In: *IEEE power and energy magazine* 8.1, pp. 18–28.

Fekri, Mohammad Navid et al. (2021). "Deep learning for load forecasting with smart meter data: Online Adaptive Recurrent Neural Network". In: *Applied Energy* 282, p. 116177.

Fernandes, Silas EN et al. (2018). "A Probabilistic Optimum-Path Forest Classifier for Non-Technical Losses Detection". In: *IEEE Transactions on Smart Grid*.

Gao, Yuanqi, Brandon Foggo, and Nanpeng Yu (2019a). "A Physically Inspired Data-Driven Model for Electricity Theft Detection With Smart Meter Data". In: *IEEE Transactions on Industrial Informatics* 15.9, pp. 5076–5088. DOI: 10.1109/TII.2019.2898171.

— (2019b). "A Physically Inspired Data-driven Model for Electricity Theft Detection with Smart Meter Data". In: *IEEE Transactions on Industrial Informatics*.

Gellings, C et al. (2011). "Estimating the costs and benefits of the smart grid". In: *EPRI2011*.

Ghosal, Amrita and Mauro Conti (2019). "Key Management Systems for Smart Grid Advanced Metering Infrastructure: A Survey". In: *IEEE Communications Surveys & Tutorials*.

Glauner, Patrick, Jorge Augusto Meira, Lautaro Dolberg, et al. (2017). "Neighborhood features help detecting non-technical losses in big data sets". In: *2016 IEEE/ACM 3rd International Conference on Big Data Computing Applications and Technologies (BDCAT)*. IEEE, pp. 253–261.

Glauner, Patrick, Jorge Augusto Meira, Petko Valtchev, et al. (2016). "The challenge of non-technical loss detection using artificial intelligence: A survey". In: *arXiv preprint arXiv:1606.00626*.

Goudarzi, Arman et al. (2022). "A Survey on IoT-Enabled Smart Grids: Emerging, Applications, Challenges, and Outlook". In: *Energies* 15.19, p. 6984.

Gov.UK (2022). *Smart Meter Statistics in Great Britain: Quarterly Report to end March 2022*. Accessed: 2022-01-22. URL: https://www.gov.uk/government/

statistics/smart-meters-in-great-britain-quarterly-update-march-2022.

Guan, Zhitao et al. (2018). "Privacy-preserving and efficient aggregation based on blockchain for power grid communications in smart communities". In: *IEEE Communications Magazine* 56.7, pp. 82–88.

Gunturi, Sravan Kumar and Dipu Sarkar (2021). "Ensemble machine learning models for the detection of energy theft". In: *Electric Power Systems Research* 192, p. 106904.

Han, Wenlin and Yang Xiao (2016). "Non-technical loss fraud in advanced metering infrastructure in smart grid". In: *International conference on cloud computing and security*. Springer, pp. 163–172.

Han, Xue et al. (2018). "Taxonomy for evaluation of distributed control strategies for distributed energy resources". In: *IEEE Transactions on Smart Grid* 9.5, pp. 5185–5195.

Hand, David J and Robert J Till (2001). "A simple generalisation of the area under the ROC curve for multiple class classification problems". In: *Machine learning* 45.2, pp. 171–186.

Hao, Jingbo and Yang Tao (2022). "Adversarial attacks on deep learning models in smart grids". In: *Energy Reports* 8, pp. 123–129.

Hao, Jinping et al. (2015). "Sparse malicious false data injection attacks and defense mechanisms in smart grids". In: *IEEE Transactions on Industrial Informatics* 11.5, pp. 1–12.

Heaton, Jeff (2016). "An empirical analysis of feature engineering for predictive modeling". In: *SoutheastCon 2016*. IEEE, pp. 1–6.

Hegazy, Hanem I et al. (2022). "Real-Time Locational Detection of Stealthy False Data Injection Attack in Smart Grid: Using Multivariate-Based Multi-Label Classification Approach". In: *Energies* 15.14, p. 5312.

Hu, Jiankun and Athanasios V Vasilakos (2016). "Energy big data analytics and security: challenges and opportunities". In: *IEEE Transactions on Smart Grid* 7.5, pp. 2423–2436.

Hu, Wenjie et al. (2020). "Understanding electricity-theft behavior via multi-source data". In: *Proceedings of The Web Conference 2020*, pp. 2264–2274.

Huang, Ling et al. (2011). "Adversarial machine learning". In: *Proceedings of the 4th ACM workshop on Security and artificial intelligence*. ACM, pp. 43–58.

IEA (2018). *Global Energy and CO2 Status Report/The latest trends in energy and emissions*. Accessed: 2022-01-06. URL: https://www.iea.org/geco/electricity/..

— (2020). *Global Energy Review 2020*. Last accessed 05-May-2020. URL: https://www.iea.org/reports/global-energy-review-2020?utm_campaign=IEA%5C%20newsletters&utm_source=SendGrid&utm_medium=Email.

Ikpehai, Augustine, Bamidele Adebisi, and Khaled Rabie (2016). "Broadband PLC for clustered advanced metering infrastructure (AMI) architecture". In: *Energies* 9.7, p. 569.

Janssens, Olivier et al. (2016). "Data-driven multivariate power curve modeling of offshore wind turbines". In: *Engineering Applications of Artificial Intelligence* 55, pp. 331–338.

Jeyaraj, Pandia Rajan et al. (2020). "Smart grid security enhancement by detection and classification of non-technical losses employing deep learning algorithm". In: *International Transactions on Electrical Energy Systems* 30.9, e12521.

Jiang, Rong et al. (2014a). "Energy-theft detection issues for advanced metering infrastructure in smart grid". In: *Tsinghua Science and Technology* 19.2, pp. 105–120.

— (2014b). "Energy-theft detection issues for advanced metering infrastructure in smart grid". In: *Tsinghua Science and Technology* 19.2, pp. 105–120.

Jindal, Anish, Bharat Bhambu, et al. (2020). "A Heuristic-Based Appliance Scheduling Scheme for Smart Homes". In: *IEEE Transactions on Industrial Informatics* 16.5, pp. 3242–3255.

Jindal, Anish, Amit Dua, et al. (2016). "Decision tree and SVM-based data analytics for theft detection in smart grid". In: *IEEE Transactions on Industrial Informatics* 12.3, pp. 1005–1016.

Jindal, Anish, Neeraj Kumar, and Mukesh Singh (2020a). "A unified framework for big data acquisition, storage, and analytics for demand response management in smart cities". In: *Future Generation Computer Systems* 108, pp. 921–934.

— (2020b). "Internet of energy-based demand response management scheme for smart homes and PHEVs using SVM". In: *Future Generation Computer Systems* 108, pp. 1058–1068.

Jindal, Anish, Angelos K. Marnerides, et al. (2019). "Identifying Security Challenges in Renewable Energy Systems: A Wind Turbine Case Study". In: *Proceedings of the Tenth ACM International Conference on Future Energy Systems*. e-Energy '19. Phoenix, AZ, USA: Association for Computing Machinery, pp. 370–372. ISBN: 9781450366717. DOI: 10.1145/3307772.3330154. URL: https://doi.org/10.1145/3307772.3330154.

Jindal, Anish, Alberto Schaeffer-Filho, et al. (2020). "Tackling energy theft in smart grids through data-driven analysis". In: *2020 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, pp. 410–414.

Jokar, Paria, Nasim Arianpoo, and Victor CM Leung (2016). "Electricity theft detection in AMI using customers' consumption patterns". In: *IEEE Transactions on Smart Grid* 7.1, pp. 216–226.

Judge, Malik Ali et al. (2022). "Overview of smart grid implementation: Frameworks, impact, performance and challenges". In: *Journal of Energy Storage* 49, p. 104056.

Jufri, Fauzan Hanif, Victor Widiputra, and Jaesung Jung (2019). "State-of-the-art review on power grid resilience to extreme weather events: Definitions, frameworks, quantitative assessment methodologies, and enhancement strategies". In: *Applied Energy* 239, pp. 1049–1065.

Júnior, Leandro Aparecido Passos et al. (2016). "Unsupervised non-technical losses identification through optimum-path forest". In: *Electric Power Systems Research* 140, pp. 413–423.

Kawann, Cornelia (2002). "Reliability of the US electric system–Recent trends and current issues". In.

Khan, Javed Ali et al. (2015). "Comparison of Requirement Prioritization Techniques to Find Best Prioritization Technique." In: *International Journal of Modern Education & Computer Science* 7.11.

Khoo, Benjamin and Ye Cheng (2011). "Using RFID for anti-theft in a Chinese electrical supply company: A cost-benefit analysis". In: *2011 Wireless Telecommunications Symposium (WTS)*. IEEE, pp. 1–6.

Khurana, Udayan, Horst Samulowitz, and Deepak Turaga (2018). "Feature engineering for predictive modeling using reinforcement learning". In: *Thirty-Second AAAI Conference on Artificial Intelligence*.

Kim, Jin Young et al. (2019). "Detection for Non-Technical Loss by Smart Energy Theft With Intermediate Monitor Meter in Smart Grid". In: *IEEE Access* 7, pp. 129043–129053.

Krawczyk, Bartosz (2016). "Learning from imbalanced data: open challenges and future directions". In: *Progress in Artificial Intelligence* 5.4, pp. 221–232.

Krebs, Brian (2012). *FBI: Smart Meter Hacks Likely to Spread*. Accessed on: October 2022". URL: `https://krebsonsecurity.com/2012/04/fbi-smart-meter-hacks-likely-to-spread/`.

Krishna, Varun Badrinath, Carl A Gunter, and William H Sanders (2018). "Evaluating detectors on optimal attack vectors that enable electricity theft and DER fraud". In: *IEEE Journal of Selected Topics in Signal Processing* 12.4, pp. 790–805.

Leahy, Kevin et al. (2019). "Issues with Data Quality for Wind Turbine Condition Monitoring and Reliability Analyses". In: *Energies* 12.2, p. 201.

Li, Changsheng and Bo Shen (2019). "Accelerating renewable energy electrification and rural economic development with an innovative business model: A case study in China". In: *Energy Policy* 127, pp. 280–286.

Li, Jinghang and Mengqi Hu (2020). "Continuous Model Adaptation Using Online Meta-Learning for Smart Grid Application". In: *IEEE Transactions on Neural Networks and Learning Systems*.

Li, Jundong et al. (2018). "Feature selection: A data perspective". In: *ACM Computing Surveys (CSUR)* 50.6, p. 94.

Li, Meng et al. (2020). "Blockchain-based anomaly detection of electricity consumption in smart grids". In: *Pattern Recognition Letters* 138, pp. 476–482.

Li, Weixian et al. (2019). "A Novel Smart Energy Theft System (SETS) for IoT based Smart Home". In: *IEEE Internet of Things Journal*.

Lighari, Sheeraz Niaz, Birgitte Bak Jensen, Asad Ali Shaikh, et al. (2014). "Attacks and their defenses for advanced metering infrastructure". In: *2014 6th International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT)*. IEEE, pp. 148–151.

Liu, Ming et al. (2020). "Deep learning detection of inaccurate smart electricity meters: a case study". In: *IEEE Industrial Electronics Magazine* 14.4, pp. 79–90.

Liu, Ting et al. (2017). "SEDEA: State estimation-based dynamic encryption and authentication in smart grid". In: *IEEE Access* 5, pp. 15682–15693.

Liu, Yang and Shiyan Hu (2015). "Cyberthreat analysis and detection for energy theft in social networking of smart homes". In: *IEEE Transactions on Computational Social Systems* 2.4, pp. 148–158.

Liu, Yao, Peng Ning, and Michael K Reiter (2011). "False data injection attacks against state estimation in electric power grids". In: *ACM Transactions on Information and System Security (TISSEC)* 14.1, p. 13.

Lore, Kin Gwn, Devu Manikantan Shila, and Lingyu Ren (2018). "Detecting Data Integrity Attacks on Correlated Solar Farms Using Multi-layer Data Driven Algorithm". In: *2018 IEEE Conference on Communications and Network Security (CNS)*. IEEE, pp. 1–9.

Lu, Lin-Yu et al. (2019). "Intrusion detection in distributed frequency control of isolated microgrids". In: *IEEE Transactions on Smart Grid* 10.6, pp. 6502–6515.

Ma, Ruofei et al. (2013). "Smart grid communication: Its challenges and opportunities". In: *IEEE transactions on Smart Grid* 4.1, pp. 36–46.

Maamar, Assia and Khelifa Benahmed (2018). "Machine learning techniques for energy theft detection in AMI". In: *Proceedings of the 2018 International Conference on Software Engineering and Information Management*. ACM, pp. 57–62.

Mahmoud et al. (2020). "Deep Learning Detection of Electricity Theft Cyber-attacks in Renewable Distributed Generation". In: *Transactions on Smart Grid*, pp. 73–102.

Marnerides, Angelos K et al. (2014). "Power consumption profiling using energy time-frequency distributions in smart grids". In: *IEEE Communications Letters* 19.1, pp. 46–49.

Massaferro, Pablo, J. Matías Di Martino, and Alicia Fernández (2022). "Fraud Detection on Power Grids While Transitioning to Smart Meters by Leveraging Multi-Resolution Consumption Data". In: *IEEE Transactions on Smart Grid* 13.3, pp. 2381–2389. DOI: 10.1109/TSG.2022.3148817.

MathWorks, Simulink (2021). "simulation and model-based design". In: *URL https://uk. mathworks. com/products/simulink. html*.

McLaughlin, Stephen, Brett Holbert, et al. (2013). "A multi-sensor energy theft detection framework for advanced metering infrastructures". In: *IEEE Journal on Selected Areas in Communications* 31.7, pp. 1319–1330.

McLaughlin, Stephen, Dmitry Podkuiko, and Patrick McDaniel (2009). "Energy theft in the advanced metering infrastructure". In: *International Workshop on Critical Information Infrastructures Security*. Springer, pp. 176–187.

Meira, Jorge Augusto et al. (2017). "Distilling provider-independent data for general detection of non-technical losses". In: *2017 IEEE Power and Energy Conference at Illinois (PECI)*. IEEE, pp. 1–5.

Messinis, George M and Nikos D Hatziargyriou (2018). "Review of non-technical loss detection methods". In: *Electric Power Systems Research* 158, pp. 250–266.

Messinis, George M, Alexandros E Rigas, and Nikos D Hatziargyriou (2019a). "A hybrid method for non-technical loss detection in smart distribution grids". In: *IEEE Transactions on Smart Grid* 10.6, pp. 6080–6091.

— (2019b). "A hybrid method for non-technical loss detection in smart distribution grids". In: *IEEE Transactions on Smart Grid* 10.6, pp. 6080–6091.

Moghe, Rohit et al. (2009). "A scoping study of electric and magnetic field energy harvesting for wireless sensor networks in power system applications". In: *2009 IEEE Energy Conversion Congress and Exposition*. IEEE, pp. 3550–3557.

Mohammad, Rashid (2018). "AMI Smart Meter Big Data Analytics for Time Series of Electricity Consumption". In: *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*. IEEE, pp. 1771–1776.

Mukherjee, Debottam, Samrat Chakraborty, and Sandip Ghosh (2022). "Deep learning-based multilabel classification for locational detection of false data injection attack in smart grids". In: *Electrical Engineering* 104.1, pp. 259–282.

Nallathambi, Selvam (2017). "Prediction of electricity consumption based on DT and RF: An application on USA country power consumption". In: *2017 IEEE International Conference on Electrical, Instrumentation and Communication Engineering (ICEICE)*. IEEE, pp. 1–7.

Ngamchuen, Suchat and Chaiyod Pirak (2013). "Smart anti-tampering algorithm design for single phase smart meter applied to AMI systems". In: *2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology*. IEEE, pp. 1–6.

Ofgem (2022). *Counter fraud for environmental and social programmes*. Accessed on: March 2020. URL: `https://www.ofgem.gov.uk/environmental-programmes/counter-fraud-environmental-and-social-programmes`.

Ogbodo, Emmanuel U, David Dorrell, and Adnan M Abu-Mahfouz (2017). "Cognitive radio based sensor network in smart grid: Architectures, applications and communication technologies". In: *IEEE Access* 5, pp. 19084–19098.

Otuoze, Abdulrahaman Okino, Mohd Wazir Mustafa, and Raja Masood Larik (2018). "Smart grids security challenges: Classification by sources of threats". In: *Journal of Electrical Systems and Information Technology* 5.3, pp. 468–483.

Ozger, Mustafa, Oktay Cetinkaya, and Ozgur B Akan (2018). "Energy harvesting cognitive radio networking for iot-enabled smart grid". In: *Mobile Networks and Applications* 23.4, pp. 956–966.

Pal, Seemita, Biplab Sikdar, and Joe H Chow (2017). "Classification and detection of PMU data manipulation attacks using transmission line parameters". In: *IEEE Transactions on Smart Grid* 9.5, pp. 5057–5066.

Pelletier, Francis, Christian Masson, and Antoine Tahan (2016). "Wind turbine power curve modelling using artificial neural network". In: *Renewable Energy* 89, pp. 207–214.

Peng, Chen et al. (2019). "A survey on security communication and control for smart grids under malicious cyber attacks". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 49.8, pp. 1554–1569.

Peng, Yanlin et al. (2021). "Electricity Theft Detection in AMI Based on Clustering and Local Outlier Factor". In: *IEEE Access* 9, pp. 107250–107259.

Peppanen, Jouni et al. (2015). "Leveraging AMI Data for Distribution System Model Calibration and Situational Awareness". In: *IEEE Transactions on Smart Grid* 6.4, pp. 2050–2059. DOI: 10.1109/TSG.2014.2385636.

Pfeifer, S., N. Fildes, and A. Ram (2018). *Energy sector on alert for cyber attacks on UK power network*. URL: https://www.ft.com/content/d2b2aaec-4252-11e8-93cf-67ac3a6482fd.

Possebon, Isadora et al. (2019). "Improved Network Traffic Classification Using Ensemble Learning". In: *IEEE Symposium on Computers and Communications (ISCC) 2019*.

Prado, Josue Campos do et al. (2019). "The Next-Generation Retail Electricity Market in the Context of Distributed Energy Resources: Vision and Integrating Framework". In: *Energies* 12.3, p. 491.

Punmiya, Rajiv and Sangho Choe (2019). "Energy Theft Detection Using Gradient Boosting Theft Detector With Feature Engineering-Based Preprocessing". In: *IEEE Transactions on Smart Grid* 10.2, pp. 2326–2329.

Qi, Junjian et al. (2016). "Cybersecurity for distributed energy resources and smart inverters". In: *IET Cyber-Physical Systems: Theory & Applications* 1.1, pp. 28–39.

Raggi, Livia et al. (2020). "Non-Technical Loss Identification by Using Data Analytics and Customer Smart Meters". In: *IEEE Transactions on Power Delivery*.

Ramos, Caio C. O. et al. (2011). "What is the importance of selecting features for non-technical losses identification?" In: *2011 IEEE International Symposium of Circuits and Systems (ISCAS)*, pp. 1045–1048. DOI: 10.1109/ISCAS.2011.5937748.

Razavi, Rouzbeh and Martin Fleury (2019). "Socio-economic predictors of electricity theft in developing countries: An Indian case study". In: *Energy for Sustainable Development* 49, pp. 1–10.

Razavi, Rouzbeh, Amin Gharipour, et al. (2019). "A practical feature-engineering framework for electricity theft detection in smart grids". In: *Applied Energy* 238, pp. 481–494.

RB, Sharma and Dhole GM (2015). "A Survey of Wide Area Measurment Technology in Electrical Networks". In: *2015 International Conference on Computing Communication Control and Automation*. IEEE, pp. 521–526.

Ren, Chao et al. (2014). "Optimal parameters selection for BP neural network based on particle swarm optimization: A case study of wind speed forecasting". In: *Knowledge-based systems* 56, pp. 226–239.

Rezaee, Mohammad and Mohammad Hossein Yaghmaee Moghaddam (2019). "SDN-based Quality of Service Networking for Wide Area Measurement System". In: *IEEE Transactions on Industrial Informatics*.

Rouzbahani, Hossein Mohammadi, Hadis Karimipour, and Lei Lei (2020). "An ensemble deep convolutional neural network model for electricity theft detection in smart grids". In: *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, pp. 3637–3642.

Saeed, Muhammad Salman et al. (2020). "Detection of non-technical losses in power utilities—A comprehensive systematic review". In: *Energies* 13.18, p. 4727.

Safe, Stay Energy (2018). *Report Energy Theft — Electricity Theft, Gas Theft in the UK*. Accessed on: March 2020. URL: https://www.stayenergysafe.co.uk/.

Salinas, Sergio, Ming Li, and Pan Li (2012). "Privacy-preserving energy theft detection in smart grids". In: *2012 9th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON)*. IEEE, pp. 605–613.

Salinas, Sergio A and Pan Li (2016). "Privacy-preserving energy theft detection in microgrids: A state estimation approach". In: *IEEE Transactions on Power Systems* 31.2, pp. 883–894.

Sanjab, Anibal et al. (2016). "Smart grid security: Threats, challenges, and solutions". In: *arXiv preprint arXiv:1606.06992*.

Schuld, Maria, Ilya Sinayskiy, and Francesco Petruccione (2015). "An introduction to quantum machine learning". In: *Contemporary Physics* 56.2, pp. 172–185.

Shaaban, Mostafa et al. (2021). "Data-Driven Detection of Electricity Theft Cyber-attacks in PV Generation". In: *IEEE Systems Journal*.

Sharma, Shalini and Angshul Majumdar (2020). "Unsupervised Detection of Non-Technical Losses via Recursive Transform Learning". In: *IEEE Transactions on Power Delivery*.

Shilay, Devu Manikantan et al. (2017). "Catching Anomalous Distributed Photovoltaics: An Edge-based Multi-modal Anomaly Detection". In: *arXiv preprint arXiv:1709.08830*.

Shokoya, NO and AK Raji (2019). "Electricity Theft: A Reason to Deploy Smart Grid in South Africa". In: *2019 International Conference on the Domestic Use of Energy (DUE)*. IEEE, pp. 96–101.

Siegel, D. (2012). *Electromagnetic Harveste*. Accessed on: March 2020. URL: `http://dennissiegel.de/works/electromagnetic-harvester/#more`.

Singh, Sandeep Kumar, Ranjan Bose, and Anupam Joshi (2019). "Energy theft detection for AMI using principal component analysis based reconstructed data". In: *IET Cyber-Physical Systems: Theory & Applications* 4.2, pp. 179–185.

Singh, Sandeep Kumar, Kush Khanna, et al. (2017). "Joint-transformation-based detection of false data injection attacks in smart grid". In: *IEEE Transactions on Industrial Informatics* 14.1, pp. 89–97.

Smith, Thomas B (2004). "Electricity theft: a comparative analysis". In: *Energy policy* 32.18, pp. 2067–2076.

Sobczak, Blake (2019). *'Cyber event' disrupted U.S. grid networks — DOE*. URL: `https://www.eenews.net/stories/1060242741`.

Solanki, Bharatkumar V, Kankar Bhattacharya, and Claudio A Canizares (2017). "A sustainable energy management system for isolated microgrids". In: *IEEE Transactions on Sustainable Energy* 8.4, pp. 1507–1517.

Staffell, Iain and Stefan Pfenninger (2018). "The increasing impact of weather on electricity supply and demand". In: *Energy* 145, pp. 65–78.

Sundararajan, Aditya et al. (2019). "Survey on synchrophasor data quality and cybersecurity challenges, and evaluation of their interdependencies". In: *Journal of Modern Power Systems and Clean Energy* 7.3, pp. 449–467.

Tajer, Ali (2017). "False data injection attacks in electricity markets by limited adversaries: stochastic robustness". In: *IEEE Transactions on Smart Grid* 10.1, pp. 128–138.

Takiddin, Abdulrahman, Muhammad Ismail, and Erchin Serpedin (2022). "Robust Data-Driven Detection of Electricity Theft Adversarial Evasion Attacks in Smart Grids". In: *IEEE Transactions on Smart Grid* 14.1, pp. 663–676.

Takiddin, Abdulrahman, Muhammad Ismail, Usman Zafar, et al. (2020). "Robust electricity theft detection against data poisoning attacks in smart grids". In: *IEEE Transactions on Smart Grid* 12.3, pp. 2675–2684.

Tan, Song et al. (2017). "Survey of security advances in smart grid: A data driven approach". In: *IEEE Communications Surveys & Tutorials* 19.1, pp. 397–422.

Tazi, Nacef and Youcef Bouzidi (2020). "Evolution of wind energy pricing policies in France: Opportunities and new challenges". In: *Energy Reports* 6, pp. 687–692.

Tian, Di-An and Giovanni Sansavini (2016). "Impact of degraded communication on interdependent power systems: the application of grid splitting". In: *Electronics* 5.3, p. 49.

Tuballa, Maria Lorena and Michael Lochinvar Abundo (2016). "A review of the development of Smart Grid technologies". In: *Renewable and Sustainable Energy Reviews* 59, pp. 710–725.

Vaughan, A (2018). "EU raises renewable energy targets to 32% by 2030". In: *The Guardian*.

Venayagamoorthy, Ganesh Kumar et al. (2016). "Dynamic energy management system for a smart microgrid". In: *IEEE transactions on neural networks and learning systems* 27.8, pp. 1643–1656.

Wang, Chengshan et al. (2017). *Smart electricity distribution networks*. CRC Press.

Wang, Yun et al. (2018). "Wind power curve modeling and wind power forecasting with inconsistent data". In: *IEEE Transactions on Sustainable Energy* 10.1, pp. 16–25.

Wen, Mi et al. (2021). "FedDetect: A Novel Privacy-Preserving Federated Learning Framework for Energy Theft Detection in Smart Grid". In: *IEEE Internet of Things Journal*.

Weslowski, J (1976). "Utilities launch assault to halt theft of power". In: *Electr. Light Power (Boston);(United States)* 54.10.

Wilcox, Tom et al. (2019). "A Big Data platform for smart meter data analytics". In: *Computers in Industry* 105, pp. 250–259.

Xia, Xiaofang et al. (2022). "Detection methods in smart meters for electricity thefts: A survey". In: *Proceedings of the IEEE* 110.2, pp. 273–319.

Xie, Le, Yilin Mo, and Bruno Sinopoli (2010). "False data injection attacks in electricity markets". In: *2010 First IEEE International Conference on Smart Grid Communications*. IEEE, pp. 226–231.

Xue, Dongbo, Xiaorong Jing, and Hongqing Liu (2019). "Detection of false data injection attacks in smart grid utilizing ELM-based OCON framework". In: *IEEE Access* 7, pp. 31762–31773.

Yan, Zhongzong and He Wen (2021). "Performance Analysis of Electricity Theft Detection for the Smart Grid: An Overview". In: *IEEE Transactions on Instrumentation and Measurement*.

Yao, Donghuan et al. (2019). "Energy Theft Detection with Energy Privacy Preservation in the Smart Grid". In: *IEEE Internet of Things Journal*.

Yip, Sook-Chin, Wooi-Nee Tan, et al. (2018). "An anomaly detection framework for identifying energy theft and defective meters in smart grids". In: *International Journal of Electrical Power & Energy Systems* 101, pp. 189–203.

Yip, Sook-Chin, KokSheik Wong, et al. (2017a). "Detection of energy theft and defective smart meters in smart grids using linear regression". In: *International Journal of Electrical Power & Energy Systems* 91, pp. 230–240.

— (2017b). "Detection of energy theft and defective smart meters in smart grids using linear regression". In: *International Journal of Electrical Power & Energy Systems* 91, pp. 230–240.

Yorukoglu, Sinan et al. (2016). "The effect of the types of network topologies on nontechnical losses in secondary electricity distribution systems". In: *IEEE Transactions on Industry Applications* 52.5, pp. 3631–3643.

Yuan, Xiaodong, Min-gming Shi, and Zhengyang Sun (2015). "Research of electricity stealing identification method for distributed PV based on the least squares approach". In: *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*. IEEE, pp. 2471–2474.

Yuan, Xiaodong, Mingming Shi, and Zhengyang Sun (2015). "Research status of electricity-stealing identification technology for distributed PV". In: *2015 5th International Conference on Electric Utility Deregulation and Restructuring and Power Technologies (DRPT)*. IEEE, pp. 2031–2034.

Zanetti, Marcelo et al. (2017). "A tunable fraud detection system for advanced metering infrastructure using short-lived patterns". In: *IEEE Transactions on Smart Grid* 10.1, pp. 830–840.

Zhang, Yang, Tao Huang, and Ettore Francesco Bompard (2018). "Big data analytics in smart grids: a review". In: *Energy Informatics* 1.1, p. 8.

Zhang, Ying, Jianhui Wang, and Bo Chen (2020). "Detecting false data injection attacks in smart grids: A semi-supervised deep learning approach". In: *IEEE Transactions on Smart Grid* 12.1, pp. 623–634.

Zhao, Bo et al. (2018). "Energy Management of Multiple Microgrids Based on a System of Systems Architecture". In: *IEEE Transactions on Power Systems* 33.6, pp. 6410–6421.

Zheng, Kedi et al. (2018). "A Novel Combined Data-Driven Approach for Electricity Theft Detection". In: *IEEE Transactions on Industrial Informatics* 15.3, pp. 1809–1819.

Zheng, Zibin et al. (2018). "Wide and deep convolutional neural networks for electricity-theft detection to secure smart grids". In: *IEEE Transactions on Industrial Informatics* 14.4, pp. 1606–1615.