Computers and Electrical Engineering

Context-aware Edge Computing and Internet of Things in Smart Grids: a systematic mapping study

--Manuscript Draft--

Manuscript Number:	COMPELECENG-D-21-01583R4
Article Type:	Research Paper
Keywords:	Smart Grid; Data Analysis; Edge computing; Internet of Things; MACHINE LEARNING; context awareness
Abstract:	Smart Grid (SG) is a relatively new subject in which intelligence comes from computer systems. Edge Computing (EC) processes data near the SG information source -- potentially attenuating latency problems. While Internet of Things (IoT) collects, transmits and stores data, EC allows the employment of data analysis with Machine Learning (ML) at its edge. This work uses a systematic mapping methodology to encompass 15081 works related to four SGs technologies: data analysis, EC, IoT, and context awareness. With 37 works after the filtering process, this review revealed that most papers use one to three approaches, while only two use all four technologies. The results also indicate that EC has been extensively used in SG solutions, with 22 selected studies. Distinctively, only 9 works use context awareness, which may indicate a path for future developments in SG. The study also allowed the learning of 7 lessons that are presented in this paper.

****

Response Letter

****

Dear Editor.

We would like to thank you and the anonymous reviewers for the opportunity of improving the paper. The study showed how Machine Learning, Internet of Things, Edge Computing, and Context-Aware computing could improve the smart grid systems.

To meet the requirements of the Computers and Electrical Engineering Journal, we made the requested changes. The abstract has up to 150 words. The conclusion has no references and is up to 300 words. We have reduced the number of references to 25 papers. Table 4 has a new column to display the DOI number of each reviewed study, allowing the reader to find any of these papers. Still, in table 4, the reference column changed to Id. The id column can guide the reader related to the reviewed papers.

The paper size has been reduced to 27 pages. We optimize the space of figures and tables to reduce the paper length as much as possible. Despite the 27 pages, the article has 23 pages, except for references. Furthermore, we assure that the authors have read the final version to improve as much as possible the quality of the English language and to ensure the quality and accuracy of technical and scientific content.

Thank you once again for the opportunity, and we are at your disposal for any clarification or further improvements.

Best regards,

The authors

Highlights

We review the state of the art related to smart grid computational techniques.

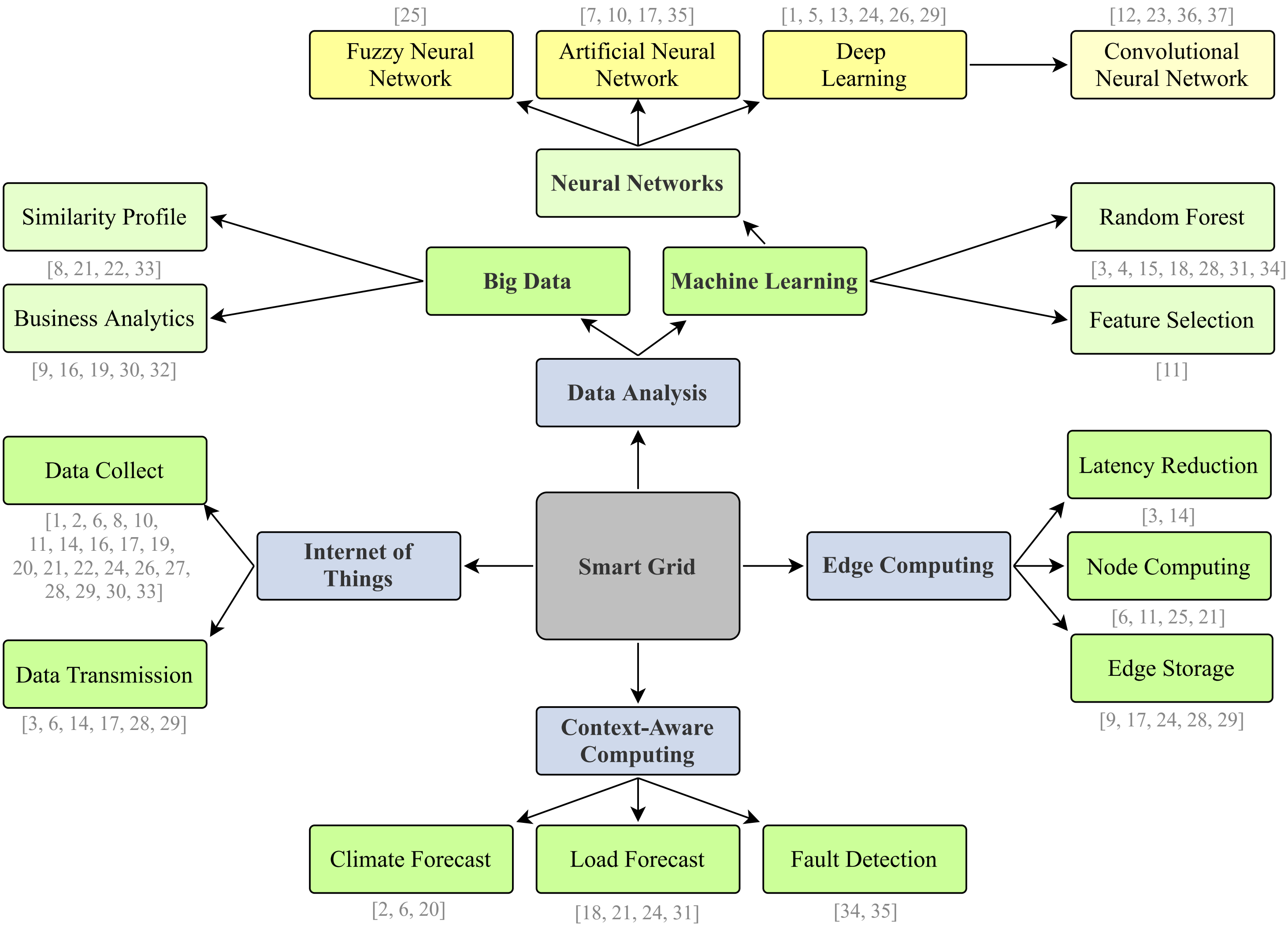
The paper covers internet of things, edge computing, and context-aware computing.

The work considers literature of the last decade through a well-defined methodology.

The study covers 15081 papers, answering one general and eight focused issues.

The paper proposes a taxonomy for smart grid Technologies.

Graphical Abstract



Context-aware Edge Computing and Internet of Things in Smart Grids: a systematic mapping study

Jorge Arthur Schneider Aranda[a,*], Ricardo dos Santos Costa[b], Vitor Werner de Vargas[a], Paulo Ricardo da Silva Pereira[b], Jorge Luis Victória Barbosa[a,b], Marcelo Pinto Vianna[c]

[a] *Applied Computing Graduate Program (PPGCA), University of Vale do Rio dos Sinos (UNISINOS), Ave. Unisinos, 950 – 93.022-750, São Leopoldo, RS, Brazil*
[b] *Electric Engineering Graduate Program (PPGEE), University of Vale do Rio dos Sinos (UNISINOS), Ave. Unisinos, 950 – 93.022-750, São Leopoldo, RS, Brazil*
[c] *CEEE-D Power Utility, St. Joaquim Pôrto Vilanova, 201 Porto Alegre, Brazil*

Abstract

Smart Grid (SG) is a relatively new subject in which intelligence comes from computer systems. Edge Computing (EC) processes data near the SG information source – potentially attenuating latency problems. While Internet of Things (IoT) collects, transmits and stores data, EC allows the employment of data analysis with Machine Learning (ML) at its edge. This work uses a systematic mapping methodology to encompass 15081 works related to four SGs technologies: data analysis, EC, IoT, and context awareness. With 37 works after the filtering process, this review revealed that most papers use one to three approaches, while only two use all four technologies. The results also indicate that EC has been extensively used in SG solutions, with 22 selected studies. Distinctively, only 9 works use context awareness, which may indicate a path for future developments in SG. The study also allowed the learning of 7 lessons that are presented in this paper.

Keywords: Smart Grid, Data Analysis, Edge Computing, Internet of Things, Machine Learning, Context Awareness

*Corresponding author
Email address: jsaranda@unisinos.br (Jorge Arthur Schneider Aranda)

1. Introduction

Electric power systems are infrastructures linked to aspects of human development and well-being. As the need for electric power keeps increasing, the development of technologies for effective and intelligent management of power grids is strategic. A Smart Grid (SG) is a type of Cyber-Physical System (CPS) that enables an electric power grid to coexist with Information and Communication Technologies (ICT)[1].

A CPS is considered intelligent, secure, reliable, and economical. However, SGs are still complex due to the sheer number of different devices (sensors, smart meters, and actuators), the extensive geographical size, and the use of computationally intensive models for control, analysis, optimization, and simulation [2]. Over the years, Cloud Computing (CC) infrastructures have been the dominant solutions to handle heavy computational tasks related to SG applications [3]. CC offers on-demand services, enabling data gathered across the grid to be processed efficiently. Therefore, the architectures of SGs and CC can be combined to achieve satisfactory results [4].

Internet and wireless communication technologies have provided a convenient channel for information exchange in people's daily life. By 2019, the number of global mobile terminals increased exponentially, reaching about 2.8 billion [5]. Especially with the advance of artificial intelligence, the number of lightweight intelligent devices has also increased exponentially. The interconnection of all things has become the primary trend of development for wireless communication networks, and the Internet [5]. The Internet of Things (IoT) era can support the SGs with a large number of sensors, actuators, and mobile devices deployed at the network edge. However, a considerable part of the computing tasks generated by these devices requires timely, and context-aware processing [6].

As a result, processing massive data traffic is a crucial feature for SG future systems. Furthermore, high data rate and low delivery latency have become two key performance indicators of the computing systems applied to SGs. These indicators imply that computation devices need powerful data processing and high data rate transmission links to transfer data traffic for the internet and wireless communication networks, respectively [7].

The development of computing architectures and paradigms to provide robust internet data processing occurred in the past few decades. The design of the computer, the internet, and information technologies have brought people into an era of information (big data). There is a need for changes in computer networks, storage, and applications to meet this amount of information requirements [8].

CC provides powerful capabilities to address these computing challenges with centralized processing and storage resources. For instance, CC provides elastic services and data-intensive analysis for end-users over a Wide Area Network (WAN). Therefore, users can be empowered with seemingly unlimited resources without building new computing infrastructures. However, although the centralization of computing resources in the cloud facilitates resource management and maintenance, it complicates the fulfillment of service demands for the new trend of delay-sensitive IoT applications, especially in SGs – which need constant monitoring in broad geographic areas [9].

The first issue is the WAN latency, which is unlikely to be improved in the foreseeable future since CC mainly focuses on improving the efficiency of bandwidth and links [5]. The second issue is the insufficiency in traffic capacity of WAN, by the increasing amount of data generated by IoT devices. Lastly, CC has intrinsic disadvantages of supporting context-aware computing in IoT applications since CC implemented remotely, through centralized computing [7].

In order to address these issues, new network paradigms emerged, offering computing resources in the proximity of end-users. In such a way, delay-sensitive and context-aware services can be offered without the involvement of WAN [10]. Emerging network computing paradigms, such as fog computing and Edge Computing (EC), employ small-scale edge servers with limited computational resources to timely serve end-users at the network edge. Fog and EC can be deemed cloud service extensions to the network edge since they exploit similar computation offloading and storage management schemes. The prediction of

SGs' next optimal step can be achieved through context-aware techniques [11], such as context histories [12] and context prediction [13].

This paper focuses on find current computing technologies used in the SG field, thus serving as a basis for future works. This work found 15081 papers related to IoT, EC, data analysis applied to the SG searching through seven electrical and computing focused digital libraries through a systematic mapping. From these, a filtering process selected 37 papers for analyses and discussions.

The paper is organized as follows. Section II describes related works, section III details the research method used in this literature review, section IV presents the results, focused on the research questions, section V discusses on the findings and, finally, section VI provides the conclusions.

2. Related Works

During the search stage, literature review articles of SGs returned by the search string were considered as related works for reviewing the literature within the same area of a domain that this work proposes. Five reviews and surveys address EC, IoT, data analysis, and context-aware concepts applied to SGs [14, 15, 16, 17, 18]. Three studies only consider the effects on the SG after an event occurs [14, 15, 17]. Using artificial intelligence techniques such as Machine Learning (ML), these three studies do not consider papers that analyze data in real-time partially meeting the requirements for data analysis. Other studies [15, 17] consider EC partially, since the review studies do not process nor analyze data at the edges, using EC only to collect and transfer data. Conversely, the present work reviewed papers that consider data analysis, IoT, EC, and context-aware techniques applied to SG.

The work of Davody and Beni [14] focuses primarily on IoT and SG, analyzing their advantages, challenges, and practical solutions. The authors address critical aspects, like big data, expenditure reduction, and system security.

Zhen et al. [15] summarized the key big data technologies and research ideas and discussed the problems faced by big data technologies in SG. The main problems refer to security and data management in SG, as well as latency in data acquisition.

Shi et al. [16] published a survey on artificial intelligence techniques applied to SG. The authors concluded that these techniques can predict SG stability. Alternatively, communication problems may occur when evaluating the research models in an SG.

The survey of Ibrahim et al. [17] demonstrates the increasing interest and expansion in the use of ML techniques in SG. According to the authors, some issues remain open and are worth further research, such as the high-performance data processing, and intelligent decision-making in large-scale complex multi-energy systems, lightweight ML, and EC.

The literature review performed by Cheng and Yu [18] focuses on introducing and summarizing seven usual ML methods in the field of SG: reinforcement learning, deep learning, transfer learning, parallel learning, hybrid learning, adversarial learning, and ensemble learning.

Furthermore, these works [14, 15, 16, 17, 18] have discussed advantages, architectures, applications, and research issues. Table 1 shows a comparison between the related works approach.

Table 1: Comparison between related works

Reference	Data Analysis	Internet of Things	Edge Computing	Context- Aware
Davody et al.[14]	Partially	Yes	No	No
Zhen et al. [15]	Partially	Yes	Partially	No
Shi et al. [16]	Yes	No	No	Yes
Ibrahim et al. [17]	Partially	Yes	Partially	No
Cheng et al. [18]	Yes	No	No	Yes

3. Research Method

This paper uses a systematic mapping study methodology for conducting a literature review of research works that investigated how IoT, EC, data analysis, and context awareness can aid in the implementation of SGs [19, 20]. The main objective of such a review is to identify evidence and trends in collections of literary works related to a topic of interest, thus reducing bias when using single references. Based on the guidelines proposed by Pinciroli et al. [21], the systematic mapping applied the following steps: Research questions; Research process; Criteria for filtering results.

3.1. Research Questions

The research questions delineated the discovery of papers related to data analysis, IoT, context awareness, and EC applied in SGs. Hence, the study defined one General question (GQ), six Focused Questions (FQ), and two Statistical Questions (SQ). Table 2 presents the questions.

The GQ sought basic information regarding the technologies used in SGs. FQs explore quantitative details of the selected papers, such as the most common data analysis techniques or how many studies use IoT. Finally, SQs aimed to verify the publications' chronological data and type of venue.

Table 2: Research Questions

Reference	**Question**
General Questions	
GQ1	How data analysis and Internet of Things been used to support Edge Computing on Smart Grids?
Focused Questions	
FQ1	Which are the data analysis techniques applied to Edge-Computing in smart grid?
FQ2	Are there studies that consider Contexts, Context Histories and Context Prediction, according to the Dey's definition [22]?
FQ3	Which are the adaptation strategies used to improve the data management in Edge-Computing applied to Smart Grid?
FQ4	How has the Internet of Things been used for Edge-Computing in Smart Grids?
FQ5	How the works used Big Data to support Edge-Computing in Smart Grids?
FQ6	How has the Machine Learning prediction's been used to support Edge Computing in Smart Grids?
Statistical Questions	
SQ1	What is the number of publications per type?
SQ2	How Many publications occurred per year?

3.2. Research Process

The study defined three steps for the research process: specify the search string, select databases, and find the results. The first step identified the main terms and their most relevant synonyms. The terms chosen were “smart grid”, “data analysis” and “edge computing” as primary terms, and “smart energy”, “big data”, “machine learning”, “deep learning”, “data analytics”, “intelligent edge computing”, “edge computing”, “intelligent edge computing”, “internet of thing” and “fog computing” as synonyms as displayed in Table 3.

Table 3: Definition of the Search String

Major Terms	Search Terms
Smart Grid	((smart grid OR smart energy) AND
Data Analysis	(data analysis, OR big data OR machine learning OR deep learning OR data analytics) AND
Edge Computing	(edge computing OR intelligent edge computing OR edge computing OR intelligent edge computing OR internet of thing OR fog computing))

The search string elaboration consists of the definition of the major terms and their synonyms. After defining the search string, the search process encompassed seven digital libraries: ACM, IEEE Xplore, Scopus, Google Scholar, Springer Link, Science Direct, and Wiley. The selection prioritized electrical and computer science databases, which had previously been used in recent systematic review studies [23, 24].

Research in ACM and IEEE Xplore, Science Direct, and Wiley required the use of an advanced search feature. Google Scholar and Scopus search required a combination of the summary and title fields in the advanced search option. Finally, in Springer Link removing documents categorized as “preview only” and select the search filter titled ‘computer science’ to obtain relevant results.

3.3. Study Filtering

After gathering literary works through the search string, the filtration process sorts the papers related to the research area and removes those that are not.

The following rejection criteria (RC) allowed to remove papers: the study must have been published in a journal, conference, or workshop (IC 1); the study must be related to the proposed theme – data analysis, EC, IoT and context awareness in SG (IC 2); the study must be a complete paper (IC 3).

The following inclusion criteria (IC) allowed the filtering of papers: studies published prior to 2010 (RC 1); studies not written in English (RC 2); studies published as dissertations or theses (RC 3); studies which did not have any relation to the research questions (RC 4).

The inclusion and rejection criteria enabled the attainment of the most relevant studies and removed any noise generated in the research. Figure 1 shows the result of the filtering process. The initial filtering of papers consisted of removing impurities that did not comply with the RC 1, 2, and 3. Lastly, the RC 4 enabled the extraction of any residuals through the three-pass method [25]. The first step of the three-pass method comprised four stages: 1) read the title, the abstract, and the introduction of each paper; 2) read the titles of each section and subsection; 3) look at the mathematical equations (if available) to review whether they are consistent with the theoretical grounds presented in the paper; 4) read the conclusions. The second step involved carefully reviewing figures, diagrams, and other illustrations of papers, with specific attention given to figures. Finally, the third step was to read the full text, observing the RC 4.

4. Results

The filtering process resulted in 37 papers. At this stage, the selected papers were analyzed according to their objectives. Table 4 presents papers indicating the paper Id, year of publication, data analysis technique, the use of context awareness, EC, IoT, H-Index, and paper DOI.

4.1. GQ1 - How have data analysis and Internet of Things been used to support Edge Computing in Smart Grids?

SG systems integrate several technologies. These technologies can be used for different purposes, such as forecasting electric power demand and predicting

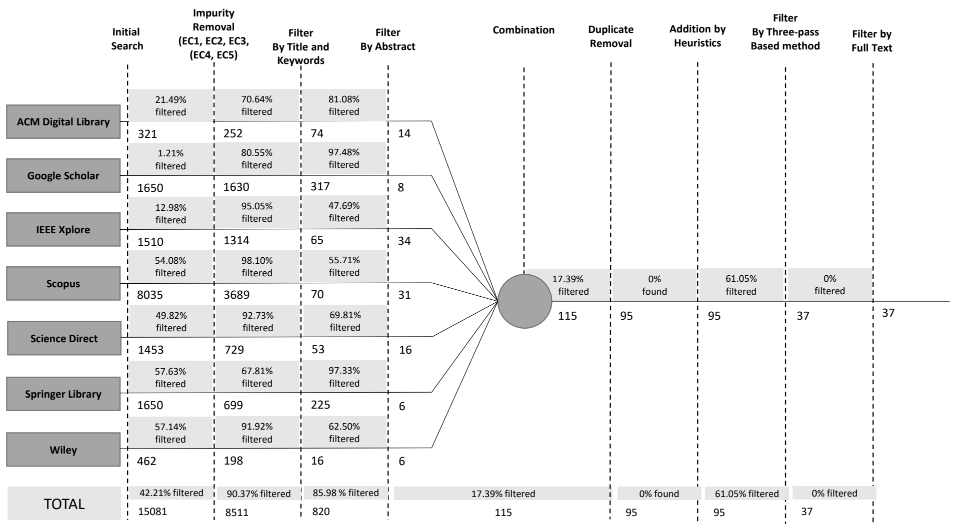


Figure 1: Process of studies filtering

problems in an electric grid.

Zhang et al. (Id 1) proposed real-time monitoring of residential load schedule. This work uses a deep learning framework through IoT devices to predict load variations throughout the day. Carvalho et al. (Id 2) presented an architecture that uses IoT devices to process collected SG data distributively. This architecture considers the network latency and sends data when the network is not busy.

The work of He et al. (Id 3) developed an ML algorithm based on causal feature selection in order to predict power events in the grid – such as outages. Liu et al. (Id 4) introduced a big data index to save SG collected data in different types of indexes, reducing the space required to save data. Omara et al. (Id 5) proposed a framework to transfer data according to the current context. In case of a sensitive event, the framework sets a high priority in the node where the event occurred.

Bin et al. (Id 6) developed a micro-service that considers changes in the network and uses this information to control company business applications. The authors' simulations show that the micro-service may reduce electrical maintenance costs.

The work proposed by Mhdawi and Al-Raweshidy (Id 7) applies a neural network (NN) framework with two approaches: firstly, the data is predicted at the edges; secondly, in case of a failure, the EC node computing tasks can be bypassed to another node. Newaz et al. (Id 8) considered an SG located on a university campus. The SG data is collected, sent to a server, and used to predict future load variations. Huang et al. (Id 9) presented a framework that monitors and reduces latency in SG networks. Using a data compression algorithm, the authors obtained 85% of reduction in latency.

The work of Raju et al. (Id 10) consists of an application that can predict the load of an SG in short periods. In order to predict these variations, the study uses different algorithms – such as Radial Basis Function (RBF), Decision Tree Regression (DTR), and random forest (RF). Elkhart and Benhlima (Id 11) showed an architecture that improves SG data management. Using big data techniques, the architecture improves the information of data analytics systems with a cost-efficiency dashboard.

The research of Tasfi et al. (Id 12) proposes a deep semi-supervised Convolutional Neural Network (CNN) with confidence sampling for electrical anomaly detection. The solution uses two sub-networks in order to achieve semi-supervised learning. While the first performs reconstruction and uses unlabelled data, the second performs classification with labeled data. Ustundagsoykan et al. (Id 13) presented a practical implementation of load forecasting with differential privacy techniques using the Tensorflow Privacy library. The authors show that data privacy guarantee can be achieved to varying degrees with a tolerable degradation in the forecasted values.

Kulkarni et al. (Id 14) developed a functional unit of the EC node, taking into account constraints – such as costs, customizations, data storage, cybersecurity, and power management. The platform was built, deployed, and applied in distributed SG applications like power quality measurements, automated metering infrastructure, and utility asset monitoring. Mousavi et al. (Id 15) proposed a decision tree-based methodology for identifying the origin of a general abnormality in SGs through data from a multi-feeder distribution system. The work of Gore et al. (Id 16) presents an IoT-based SG analyzer that efficiently utilizes the advantages put forward by IoT technology to improve situational awareness of power grids. The mobile-based software tool enables power system experts, including operators, to make decisions based on the current SG condition.

The study of Chen et al. (Id 17) introduces an EC system for IoT-based SGs to overcome the drawbacks in the current CC paradigm. Additionally, the work implements a privacy protection strategy via EC and data prediction. Vantuch et al. (Id 18) developed a boosting model to predict short-term and long-term loads. After comparing several computational models on three different regression-based criteria, the results revealed that the model outperformed its competitors in most of these comparisons.

Dalcekovic et al. (Id 19) proposed a general approach to service design considerations based on big data platforms. The method, implemented using Apache HBase, is applied in the context of demand response along with Distributed Management System (DMS) applications for managing SGs. Hasnat et al. (Id 20) designed a framework that operates providing Distributed State Estimation (DSE) locally – at the edge nodes. Focusing on Phase Measurement Units (PMUs) as an example of the industrial IoT in SGs, the framework uses an EC platform architecture to enable data analytics for DSE using the PMUs time-series measurements.

Wang et al. (Id 21) presented a data analysis and application framework for intelligent meter data based on cloud-fog computing and data contextualization. A data contextualization model based on three-dimensional (3D) maturity analysis of industrial power users is proposed to evaluate load characteristics of users from consumption behavior. Xia et al. (Id 22) developed an algorithm of distributed processing to detect malicious use of energy in SG areas. The authors claim that the work results show that this model allows power grid operators to understand quickly and intuitively the load behavior pattern and power demand of industrial power users.

In order to predict events in SG, Liu et al. (Id 23) developed a CNN using Long Short-Term Memory (LSTM). The work claims that the method increases the training speed by 61.7%, reduces Root Mean Square Error (RMSE) by 32.9%, and enhances the prediction accuracy by 1.4% compared to similar solutions. Using an MLSTM model, Alazab et al. (Id 24) considered a NN to predict the stability of an SG. According to the authors, the experimental results prove that the MLSTM approach outperforms the other ML approaches.

The work of Jurado et al. (Id 25) uses fuzzy reasoning to estimate lost data during data collection improving the accuracy by around 31.5% for different data sets tested by the authors. Mukherjee et al. (Id 26) implemented standard regression and ML-based architectures for SG load analysis and forecasting. The proposed approach predicts 97% of registers when 73% of training data have missing values.

Rodrigues et al. (Id 27) presented a low-cost smart meter methodology. According to the authors, the solution has a good potential since it has a low-cost implementation. The work of Rabie et al. (Id 28) proposes a fog data forecasting approach using EC nodes in an SG. These nodes use big data to increase the system's prediction accuracy. Additionally, Rabie et al. (Id 29) designed a methodology using outlier rejection to improve the accuracy of big data in SG. The work of Qureshi et al. (Id 30) also uses big data techniques to improve the performance of the SG network by reducing latency and package loss.

Ahmad et al. (Id 31) developed an ML model to detect problems in an SG according to the current weather. After comparing with four other ML models, the authors concluded that the proposed solution has better results. Mihailescu et al. (Id 32) provided a computational characterization system in terms of complexity, as well as an empirical analysis against real consumption data sets. Based on the macro-model of the Australian energy market, the results show a performance improvement of about 17%. Chen et al. (Id 33) created an anomaly detection monitoring consumption through IoT. The authors used profile similarity analysis to detect a possible fault in the SG network.

Zainab et al. (Id 34) developed a technique that improves the speed and the accuracy of different models for short-term prediction of SGs. According to the authors, the random tree algorithm obtained the best results using an SG dataset. The work of Qadir et al. (Id 35) develops an Artificial Neural Network (ANN) to forecast possible energy generation by solar panels. According to the authors, the prediction applying linear regression has 95% accuracy. The model of Krč et al. (Id 36) applies CNN to classify power demand for 42 different cities obtaining an average accuracy of 96%. Finally, Aldegheishem et al. (Id 37) proposed a model that combines support vector machine with CNN to detect possible outliers and electricity theft in SGs.

4.2. FQ1 - Which are the data analysis techniques applied to Edge Computing in Smart Grids?

Figure 2 shows the most used data analysis techniques. Seven works use ML-based approaches with RF (Ids 3, 4, 15, 18, 28, 31, 34). These studies use a variation of the RF algorithm to predict long-term loads, events like faults, and even weather conditions that affect the SG.

Fifteen works consider NN for ML-based solutions. Some of these studies focus on event, stability predictions, and short-term load prediction in SG (Ids 7, 10, 17, 35, 1, 5, 13, 24, 26, 29). Others studies detect anomalies and predict power demand in SG (Ids 12, 23, 36, 37). Finally, the last NN study applies fuzzy reasoning to attenuate data loss during collection, and transmission (Id 25).

Five papers use business analytics to send more information to the SG systems' management (Ids 9, 16, 19, 32, 30). Four similarity profile studies identify consumption patterns of residential neighbors or university campuses in order to predict future power demand (Ids 8, 21, 22, 33) . Finally, one study uses feature selection to provide relevant information to SG management systems (Id 11).

Table 4: Reviewed Studies

Id	Year	Data Analysis	Context Awareness	EC	IoT	H-Index	Doi
1	2019	Deep Learning	No	Yes	Yes	8	10.1145/3302505.3310069
2	2017	No	Yes	Yes	Yes	21	10.1145/3147234.3148105
3	2019	Random Forest	No	No	No	12	10.1145/3341162.3349333
4	2014	Random Forest	No	No	No	116	10.14778/2733004.2733021
5	2018	Deep Learning	No	No	No	7	10.1145/3265863.3265883
6	2019	No	Yes	Yes	Yes	29	10.1145/3358528.3358576
7	2020	ANN	No	No	No	66	10.1109/JSYST.2019.2921867
8	2014	Similarity Profile Analysis	No	Yes	Yes	16	10.1109/ICTC.2014.6983110
9	2018	Business Analytics	No	Yes	Yes	-	10.1109/icii.2018.00019
10	2020	ANN	No	Yes	Yes	14	10.1109/icosec49089.2020.9215329
11	2016	Feature Selection	No	Yes	Yes	-	10.1109/irsec.2016.7983902
12	2017	CNN	No	No	No	17	10.1109/ithings-greencom-cpscom-smartdata.2017.158
13	2019	Deep Learning	No	No	No	-	10.1109/gcwkshps45667.2019.9024520
14	2019	No	Partially	Yes	Yes	67	10.1109/jiot.2019.2898837
15	2018	Random Forest	No	No	No	-	10.1109/tdc.2018.8440570.
16	2019	Business Analytics	No	Yes	Yes	-	10.1109/i-pact44901.2019.8960098
17	2019	ANN	No	Yes	Yes	86	10.1109/access.2019.2920488
18	2018	Random Forest	No	No	No	14	10.1109/wf-iot.2018.8355123
19	2019	Business Analytics	No	Yes	Yes	-	10.1109/isncc.2017.8072030
20	2019	No	Yes	Yes	Yes	-	10.1109/gcwkshps45667.2019.9024632
21	2020	Similarity Profile Analysis	Yes	Yes	Yes	86	10.1109/ACCESS.2020.2965543
22	2018	Similarity Profile Analysis	No	No	No	86	10.1016/j.cose.2018.05.004
23	2020	CNN	No	Yes	Yes	43	10.1016/j.scs.2020.102363
24	2020	Deep Learning	Yes	Yes	No	86	10.1016/j.asoc.2016.11.040
25	2017	Fuzzy Neural Network	No	No	No	124	10.1016/j.asoc.2016.11.040
26	2020	Deep Learning	No	Yes	Yes	20	10.1016/j.suscom.2019.100356
27	2019	No	Partially	No	Yes	81	10.1016/j.measurement.2019.106890
28	2019	Random Forest	No	Yes	Yes	41	10.1007/s10586-018-2848-x
29	2020	Deep Learning	No	Yes	Yes	41	10.1007/s10586-019-02942-0
30	2019	Business Analytics	No	Yes	Yes	54	10.1007/s11277-018-5936-6
31	2020	Random Forest	Yes	Yes	No	173	10.1016/j.energy.2020.117283
32	2016	Business Analytics	No	No	No	49	10.1111/coin.12093
33	2018	Similarity ProfileAnalysis	No	Yes	Yes	63	10.1002/cpe.4737
34	2021	RandomForest	Yes	No	No	127	10.1109/access.2021.3059730
35	2021	ANN	Yes	No	No	33	10.1016/j.egyr.2021.01.018
36	2021	CNN	No	No	No	85	10.3390/su13052954
37	2021	CNN	No	No	No	127	10.1109/access.2021.3056566

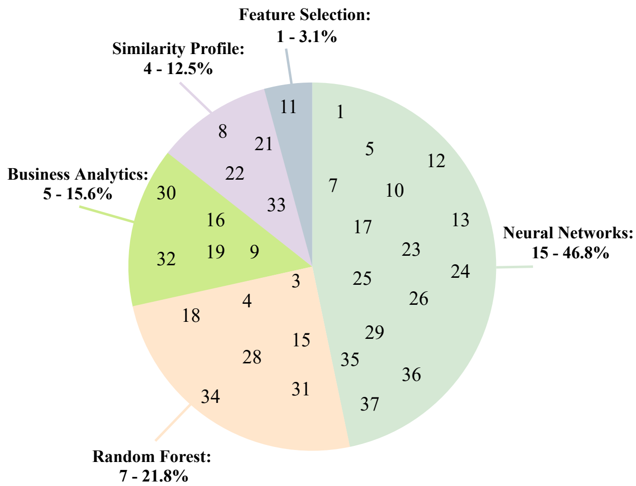


Figure 2: Types of data analysis techniques.

4.3. FQ2 - Are there any studies which consider contexts, context histories and context prediction, according to the Dey's definition [22]?

Six works consider context awareness in data analysis (Ids 2, 6, 20, 24, 21, 31). Five of them (Ids 2, 6, 20, 24, 21) use the SG context in real-time, according to voltage measurements. Four studies uses context prediction (Ids 19, 31, 34, 35) through weather forecasting in order to predict possible instability in the grid.

The remaining SG works apply CC in the network's architecture. This architecture implies that the intelligence of SG occurs far from the substations. The application of EC in addition to context-aware techniques, may help the SG systems to explore this gap.

4.4. FQ3 - Which are the adaptation strategies used for improving data management in Edge Computing applied to Smart Grids?

Five works use adaptation strategies (Ids 2, 6, 14, 20, 27). Three of these works (Ids 2, 27, 20) consist of changing the edge node to a different working node in case of a node failure. Two of the works (Ids 6, 14) adapt when the edge nodes send data to the cloud. If the edge network latency is too high in these cases, the node waits and tries to send again when the edge network has a better latency.

The use of context-aware techniques can achieve the systems' adaptivity, since systems need to know their possible next states. Context awareness may help in this case.

4.5. FQ4 - How has the Internet of Things been used for Edge Computing in Smart Grids?

IoT has been used in four different ways by 20 of the reviewed works. The first type considers data collection, storage and analysis in the IoT layer (Ids 9, 23). The second type regards data collection, and data analysis by the IoT layer (Ids 1, 16, 26, 30). The third type contemplates works which only collect data and store (temporarily or not) in the IoT layer (Ids 6, 14, 17, 28, 29). Finally, the last type considers works which only collect data through the IoT layer (Ids 2, 8, 10, 11, 19, 20, 21, 27, 33). Figure 3 shows the different uses of IoT in SGs.

4.6. FQ5 - How have the works used big data for supporting Edge Computing in Smart Grids?

Nine works use big data techniques in order to clean raw SG data using tools such as Apache Hadoop, MongoDB, Hive, Map Reduce, and Tableau (Ids 9, 11, 19, 32, 30, 8, 21, 22, 33). Additionally, the data is stored in relational or non-relational databases, where the developed solutions use the information in management dashboards, predict energy demand and determine possible anomalies in the SG.

Three works use MongoDB as non-relational databases (Ids 19, 8, 33). Other three use map-reduce techniques and store data in relational databases (Ids 16, 30, 21). Two works use Apache Hadoop (Ids 9, 32) to store SG data in clusters spread across different servers to reduce the processing required to analyze SG data. Finally, one work (Id 22) uses business intelligence and Tableau to display

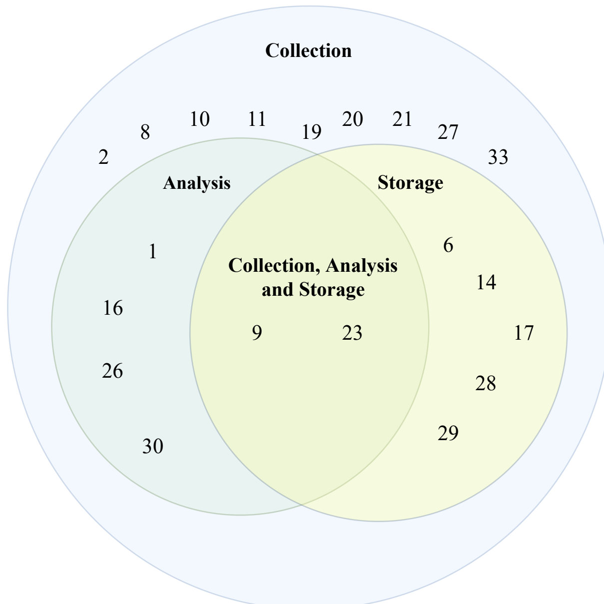


Figure 3: Different uses of Internet of Things in Smart Grid studies.

SG data from a relational database. Table 5 summarizes the big data tools and types of databases used in the reviewed papers.

Table 5: Big data tools and types of databases

Authors and Paper Id	**Big Data Tool/ Big Data Technique**	**Type of Database**
Qureshi et al. (Id 30)	Map Reduce	Relational
Gore et al. (Id 16)	Map Reduce	Relational
Wang et al. (Id 21)	Map Reduce	Relational
Mihailescu et al. (Id 32)	Apache Hadoop	Hybrid
Huang et al. (Id 9)	Apache Hadoop	Hybrid
Newaz et al. (Id 8)	Mongo DB	non-Relational
Dalcekovic (Id 19)	Mongo DB	non-Relational
Chen et al. (Id 17)	Mongo DB	non-Relational
Xia et al. (Id 22)	Tableau	Relational

4.7. FQ6 - How has Machine Learning been used for supporting Edge Computing in Smart Grids?

Four works use ML, and EC techniques simultaneously (Ids 10, 26, 28, 29). Raju et al. (Id 10) used edge nodes to clean the data. Differently, Mukherjee et al. (Id 26) applied edge nodes to analyze part of the collected data in order to help the server-side to predict futures loads. Two works of the authors Rabie et al. (Id 56, 57) temporarily store collected data in the edge nodes and send them when the SG network has low latency. The work of Krč et al. (Id 36) uses ML nodes in different cities working as edge nodes of an SG. The nodes are implemented in each city's electrical substations and perform energy demand prediction of the SG.

4.8. SQ1 - What is the number of publications per type?

Figure 4 shows the selected papers' publication data by type of venue, quantity, year, and digital library. This selection shows that 20 journal publications correspond to 54.05%, 15 conference papers account for 40.54%, and two workshop publications correspond to 5.40% of the studies reviewed in this article.

4.9. SQ2 - How many publications occurred per year?

Figure 4 shows the distribution of papers by the year of publication. These papers were analyzed from 2010 up to April 2021 since the study was performed by the end of May 2021. Research in the area has been ongoing since 2014, with an increase in 2017 and a slight variation of papers published after 2019.

5. Discussion

The studies reviewed in this paper include different types of technologies from the SG domain, data analysis, context awareness, IoT and EC. Thirteen studies apply EC, IoT, and data analysis (Ids 1, 8, 9, 10, 11, 16, 17, 19, 26, 28, 29, 30, 33). These studies consist of a node network that sends data to a server or analyses the data in the EC layer. Five papers employ the IoT, EC,

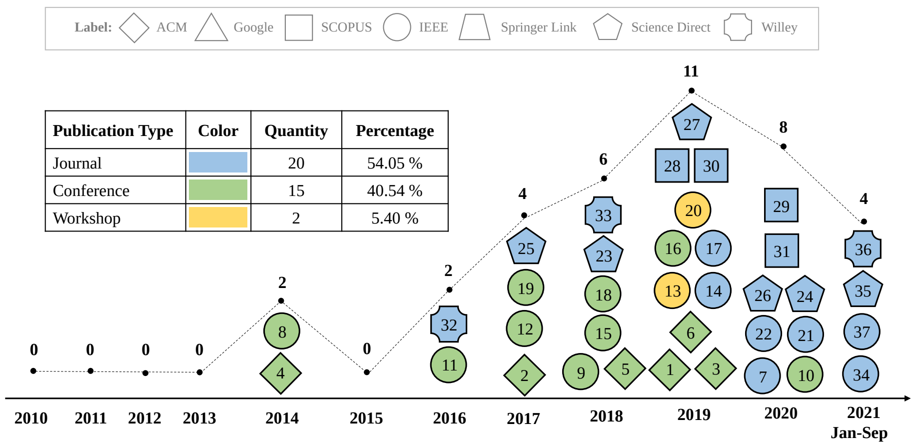


Figure 4: Publications per year by type and digital library.

and context awareness techniques (Ids 6, 14, 20, 27, 35). These papers consider network adaptation by sending data according to the latency of the SG system or SG and weather contexts to predict events that can affect the SG.

Two papers implement only context-aware EC and data analysis (Ids 24, 31). The aforementioned studies have a processing device at the edge to detect SG problems and distribute the processing in a different SG node. Pure data analysis technologies are used in eleven papers, applying ML techniques with measured and simulated data sets (Ids 3, 4, 5, 7, 12, 13, 15, 18, 32, 25, 22, 36, 37).

Finally, two works applied the four technologies extensions within the SG domain (data analysis, EC, IoT, and context awareness) (Id 2, 21). Figure 5 shows the intersections between these technologies, indicating that the majority of the SG reviewed works use data analysis, EC, and IoT at some level – although context-aware solutions are found in nine of the mapped works (Ids 2, 6, 20, 24, 21, 31, 18, 34, 35). The lack of context-aware researches may denote an opportunity for future works in the area.

IoT and data analysis applied to SGs have more works than EC and context-aware computing. Context-aware solutions have fewer studies, providing potential opportunities, specifically through works that apply weather forecasting to detect grid failures.

Only two works reducing latency between the SG edges and the operation center apply EC techniques (Ids 3, 14). Two other works applied to SG use Fuzzy NN (Id 25), and feature selection (Id 11), both ML techniques. Further studies of these techniques may point to new opportunities for future works.

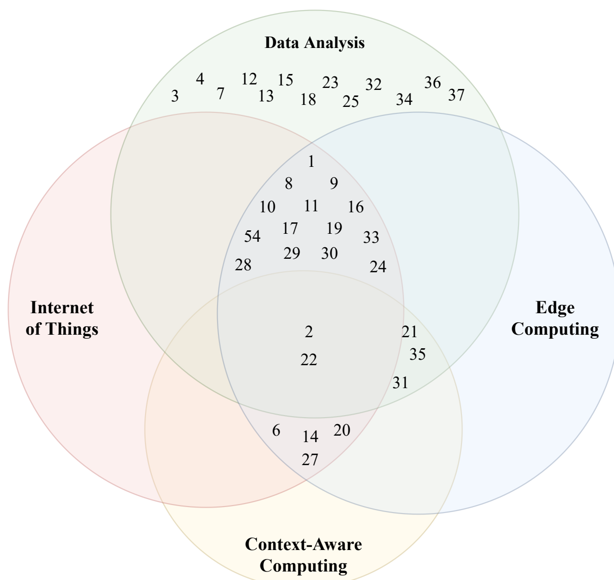


Figure 5: Intersection between Smart Grid technologies.

Considering the works analyzed in this paper, an SG taxonomy can provide an answer to the opening question in the form of a "categorized list of SG concepts". The list contains the reviewed papers' concepts related to SG, splitting them into a set of meaningful categories. The SG elements are distinguished according to different types: data analysis, EC, IoT, and context awareness.

Figure 6 displays the SG taxonomy according to the concepts' relations – with the numbers representing the reference of the selected papers. The figure shows that data analysis techniques are organized into ML (RF, feature selection, NN) and big data (business analytics, similarity profile). The NN class is divided into three other subclasses: deep learning, ANN, and fuzzy NN. The Deep learning class has a subclass called CNN. Most data analysis works are categorized as ML, indicating that artificial intelligence in SG is a well-established concept. The primary use of IoT is data collection followed by data transmission, probably due to the small sensor devices close to the data origin.

Additionally, EC is used for edge storage, node computing, and latency reduction, analyzing the data collected at the origin. This collection and analysis implicate a possible decrease in latency since less data needs to reach a data center for further analysis. Finally, few works consider context-aware computing (climate forecast, load forecast, fault detection) in the SG, which may indicate possible trends in future works.

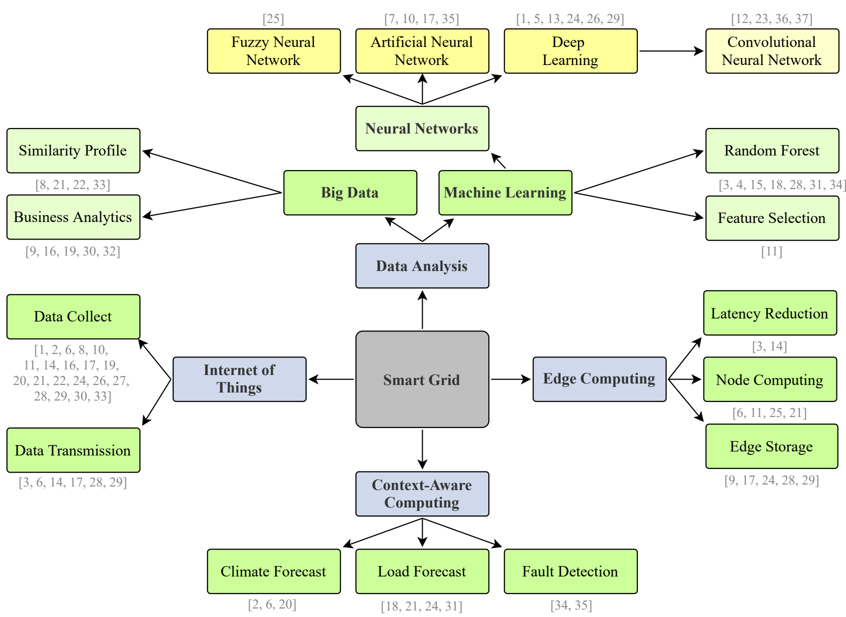


Figure 6: Taxonomy of Smart Grid technologies.

The systematic mapping of the papers provided insights, detailed solutions, and topics that need future exploration in the SG domain. In addition, lessons learned indicate concepts already established within the domain of SG. Table 6 presents seven lessons learned, summarizing contributions and observations of this study.

Table 6: Lessons learned

Lesson	Description
1	ML is a concept already established in SG applied to 62% of the works. NN correspond to 40% of the filtered works, being the most used ML technique.
2	ML combined with EC helps reducing the amount of analyzed data in the server.
3	Context histories works mainly consider weather forecasts to predict grid instability caused by bad weather.
4	Context-aware computing is only considered in 9 works (24.3%), which may denote an opportunity to explore in the SG domain.
5	EC techniques can reduce the latency between the grid edge and the grid operation center.
6	IoT is a concept already established in SG applied to 56% of the filtered works. The majority of works (54%) use sensors to perform data collection. One work (2%) applies IoT to perform only data transmission.
7	Data loss is a threat considered by five studies (13.5%). Five works propose edge storage in SG to attenuate this threat.

6. Conclusion

Research studies mainly focus on using ML-based techniques to predict energy consumption and load variations, representing 62% of the works. Additionally, 20 papers (54%) apply IoT technology to collect, store and analyze data at the edge layer. Since SGs usually work in broad geographical areas, EC is an alternative to centralized architectures, reducing delays and latency.

Only two studies (Ids 3, 21) representing 5.40% use data analysis, EC, IoT, and context awareness simultaneously. The use of these techniques enables event prediction and adaptation of SGs with more precision. The adaptation also occurs when the SG network has high latency. In these cases, the EC nodes consider the network stability to send data neatly, working as a distributed system. Furthermore, nodes can assume other node's jobs in case of failure, improving network stability and data collection.

The improvement of data collection and transmission may be need to avoid data loss, as data loss can pose a risk to the SG management. Five studies (13.5%) (Ids 9, 17, 24, 28, 29) show that data could be store at the edge of the system and retrieve when necessary, reducing the risk of data loss.

Only six works (18.7%) (Ids 2, 6, 20, 24, 21, 31) use adaptive techniques, so there is potential to explore this area. Context-aware computing may improve the use of adaptions since the SG system can optimize itself according to the context of the collected data. Future works may explore the usage of context-aware computing with SGs. In addition, the use of Context Histories to organize the data will allow pattern analysis, context prediction and similarity analysis. These strategies for handling context histories will improve the analysis of the data, mainly allowing the prediction and recommendation oriented to SG.

Acknowledgment

This work was financed in part by CNPq/Brazil (National Council for Scientific and Technological Development - http://www.cnpq.br), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001 and CEEE Equatorial (https://ceee.equatorialenergia.com.br/). We would also like to thank the Unisinos University (http://www.unisinos.br/global/en/)

References

[1] M. Z. Gunduz, R. Das, Cyber-security on smart grid: Threats and potential solutions, Computer Networks 169 (2020) 107094. doi:10.1016/j.comnet.2019.107094.
URL https://doi.org/10.1016/j.comnet.2019.107094

[2] M. B. Mollah, J. Zhao, D. Niyato, K.-Y. Lam, X. Zhang, A. M. Y. M. Ghias, L. H. Koh, L. Yang, Blockchain for future smart grid: A comprehensive survey 8 (1) (2021) 18–43. doi:10.1109/jiot.2020.2993601.
URL https://doi.org/10.1109/jiot.2020.2993601

[3] S. M. Putri, D. Maizana, Optimal smart grid management system in campus building 9 (3) (2020). doi:10.25077/jnte.v9n3.757.2020.
URL https://doi.org/10.25077/jnte.v9n3.757.2020

[4] G. Dileep, A survey on smart grid technologies and applications 146 (2020) 2589–2625. doi:10.1016/j.renene.2019.08.092.
URL https://doi.org/10.1016/j.renene.2019.08.092

[5] X. Deng, P. Guan, C. Hei, F. Li, J. Liu, N. Xiong, An intelligent resource allocation scheme in energy harvesting cognitive wireless sensor networks 8 (2) (2021) 1900–1912. doi:10.1109/tnse.2021.3076485.
URL https://doi.org/10.1109/tnse.2021.3076485

[6] M. Kaselimi, N. Doulamis, A. Voulodimos, E. Protopapadakis, A. Doulamis, Context aware energy disaggregation using adaptive bidirectional LSTM models, IEEE Transactions on Smart Grid 11 (4) (2020) 3054–3067. doi:10.1109/tsg.2020.2974347.
URL https://doi.org/10.1109/tsg.2020.2974347

[7] H. Xu, Fastest adaptive estimation algorithms for topological structure errors in smart grid networks, Computer Communications 160 (2020) 197–203. doi:10.1016/j.comcom.2020.05.042.
URL https://doi.org/10.1016/j.comcom.2020.05.042

[8] R. S. Gargees, G. J. Scott, Multi-stage distributed computing for big data: Evaluating connective topologies, in: 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), IEEE, 2020. doi:10.1109/ccwc47524.2020.9031227.
URL https://doi.org/10.1109/ccwc47524.2020.9031227

[9] S. A. Hashmi, C. F. Ali, S. Zafar, Internet of things and cloud computing-based energy management system for demand side management in smart grid, International Journal of Energy Research 45 (1) (2020) 1007–1022. doi:10.1002/er.6141.
URL https://doi.org/10.1002/er.6141

[10] C. Jiang, T. Fan, H. Gao, W. Shi, L. Liu, C. Cérin, J. Wan, Energy aware edge computing: A survey, Computer Communications 151 (2020) 556–580. doi:10.1016/j.comcom.2020.01.004.
URL https://doi.org/10.1016/j.comcom.2020.01.004

[11] J. A. S. Aranda, R. S. Bavaresco, J. V. de Carvalho, A. C. Yamin, M. C. Tavares, J. L. V. Barbosa, A computational model for adaptive recording of vital signs through context histories, Journal of Ambient Intelligence and Humanized Computing (2021)doi:10.1007/s12652-021-03126-8.
URL https://doi.org/10.1007/s12652-021-03126-8

[12] A. S. Filippetto, R. Lima, J. L. V. Barbosa, A risk prediction model for software project management based on similarity analysis of context histories, Information and Software Technology 131 (2021) 106497. doi:10.1016/j.infsof.2020.106497.
URL https://doi.org/10.1016/j.infsof.2020.106497

[13] J. Yu, M. Zhou, X. Wang, G. Pu, C. Cheng, B. Chen, A dynamic and static context-aware attention network for trajectory prediction 10 (5) (2021) 336. doi:10.3390/ijgi10050336.
URL https://doi.org/10.3390/ijgi10050336

[14] Z. Davoody-Beni, N. Sheini-Shahvand, H. Shahinzadeh, M. Moazzami, M. Shaneh, G. B. Gharehpetian, Application of IoT in smart grid: Challenges and solutions, in: 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), IEEE, 2019. doi:10.1109/icspis48872.2019.9066025.
URL https://doi.org/10.1109/icspis48872.2019.9066025

[15] J. Zhen, N. Qiang, L. Liu, Q. Hou, Discussion about recent applications of power big data in smart grid, in: 2019 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), IEEE, 2019. doi:10.1109/appeec45492.2019.8994522.
URL https://doi.org/10.1109/appeec45492.2019.8994522

[16] Z. Shi, W. Yao, Z. Li, L. Zeng, Y. Zhao, R. Zhang, Y. Tang, J. Wen, Artificial intelligence techniques for stability analysis and control in smart grids: Methodologies, applications, challenges and future directions, Applied Energy 278 (2020) 115733. doi:10.1016/j.apenergy.2020.115733.
URL https://doi.org/10.1016/j.apenergy.2020.115733

[17] M. S. Ibrahim, W. Dong, Q. Yang, Machine learning driven smart electric power systems: Current trends and new perspectives, Applied Energy 272 (2020) 115237. doi:10.1016/j.apenergy.2020.115237.
URL https://doi.org/10.1016/j.apenergy.2020.115237

[18] L. Cheng, T. Yu, A new generation of AI: A review and perspective on machine learning technologies applied to smart energy and electric power systems, International Journal of Energy Research 43 (6) (2019) 1928–1973. doi:10.1002/er.4333.
URL https://doi.org/10.1002/er.4333

[19] K. L. James, N. P. Randall, N. R. Haddaway, A methodology for systematic mapping in environmental sciences, Environmental Evidence (2016) 5 (1). doi:10.1186/s13750-016-0059-6.
URL https://doi.org/10.1186/s13750-016-0059-6

[20] I. D. Cooper, What is a mapping study ? , Journal of Medicine Library Association 104 (January) (2016) 76–78. doi:10.3163/1536-5050.104.1.013.
URL https://dx.doi.org/10.3163%2F1536-5050.104.1.013

[21] F. Pinciroli, J. L. B. Justo, R. Forradellas, Systematic mapping study: On the coverage of aspect-oriented methodologies for the early phases of the

software development life cycle, Journal of King Saud University - Computer and Information Sciences (2020)doi:10.1016/j.jksuci.2020.10.029.
URL https://doi.org/10.1016/j.jksuci.2020.10.029

[22] A. K. Dey, G. D. Abowd, D. Salber, A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications, Human Computer Interaction 16 (2-4) (2001) 97–166. doi:10.1207/s15327051hci16234_02.
URL https://doi.org/10.1207/S15327051HCI16234_02

[23] J. A. S. Aranda, L. P. S. Dias, J. L. V. Barbosa, J. V. de Carvalho, J. E. da Rosa Tavares, M. C. Tavares, Collection and analysis of physiological data in smart environments: a systematic mapping, (Journal of Ambient Intelligence and Humanized Computing 11 (7) (2019) 2883–2897. doi:10.1007/s12652-019-01409-9.
URL https://doi.org/10.1007/s12652-019-01409-9

[24] R. Bavaresco, D. Silveira, E. Reis, J. Barbosa, R. Righi, C. Costa, R. Antunes, M. Gomes, C. Gatti, M. Vanzin, S. C. Junior, E. Silva, C. Moreira, Conversational agents in business: A systematic literature review and future research directions, Computer Science Review 36 (2020) 100239. doi:10.1016/j.cosrev.2020.100239.
URL https://doi.org/10.1016/j.cosrev.2020.100239

[25] S. Keshav, S., How to read a paper, ACM SIGCOMM Computer Communication Review 37 (3) (2016) 83. doi:10.1145/1273445.1273458.
URL http://dx.doi.org/10.1145/1273445.1273458