Energy Informatics

REVIEW

Open Access

Enhancing resilience in complex energy systems through real-time anomaly detection: a systematic literature review

Ali Aghazadeh Ardebili[1,5*†], Oussama Hasidi[3*†], Ahmed Bendaouia[3,6*†], Adem Khalil[1†], Sabri Khalil[1†], Dalila Luceri[1†], Antonella Longo[1,2*†], El Hassan Abdelwahed[3*†], Sara Qassimi[4†] and Antonio Ficarella[1†]

†Ali Aghazadeh Ardebili, Oussama Hasidi, Ahmed Bendaouia, Adem Khalil, Sabri Khalil, Dalila Luceri, Antonella Longo, El Hassan Abdelwahed, Sara Qassimi and Antonio Ficarella have contributed equally to this work.

*Correspondence: ali.a.ardebili@unisalento.it; oussama.hasidi@ced.uca.ma; ahmed.bendaouia@edu.uca.ac.ma; antonella.longo@unisalento.it; abdewahed@uca.ac.ma

[1] Department of Engineering for Innovation, University of Salento, Lecce, Italy
[2] Italian Research Center on High Performance Computing, Big Data and Quantum Computing (ICSC), Bologna, Italy
[3] Computer Systems Engineering Laboratory (LISI), Computer Science Department, Faculty of Sciences Semlalia, Cadi Ayyad University, Marrakesh, Morocco
[4] Computer and Systems Engineering Laboratory (L2IS), Computer Science Department, Faculty of Science and Technology, Cadi Ayyad University, Marrakesh, Morocco
[5] Department of Research and Development, HSPI SpA-Roma, Rome, Italy
[6] Department of Embedded Systems and Artificial Intelligence, MAScIR, UM6P, Rabat, Morocco

Abstract

As real-time data sources expand, the need for detecting anomalies in streaming data becomes increasingly critical for cutting edge data-driven applications. Real-time anomaly detection faces various challenges, requiring automated systems that adapt continuously to evolving data patterns due to the impracticality of human intervention. This study focuses on energy systems (ES), critical infrastructures vulnerable to disruptions from natural disasters, cyber attacks, equipment failures, or human errors, leading to power outages, financial losses, and risks to other sectors. Early anomaly detection ensures energy supply continuity, minimizing disruption impacts, an enhancing system resilience against cyber threats. A systematic literature review (SLR) is conducted to answer 5 essential research questions in anomaly detection due to the lack of standardized knowledge and the rapid evolution of emerging technologies replacing conventional methods. A detailed review of selected literature, extracting insights and synthesizing results has been conducted in order to explore anomaly types that can be detected using Machine Learning algorithms in the scope of Energy Systems, the factors influencing this detection success, the deployment algorithms and security measurement to take in to consideration. This paper provides a comprehensive review and listing of advanced machine learning models, methods to enhance detection performance, methodologies, tools, and enabling technologies for real-time implementation. Furthermore, the study outlines future research directions to improve anomaly detection in smart energy systems.

Keywords: Anomaly detection, Smart energy, Cyber-physical systems, Cyber-physical security, Machine learning application, Anomalies, Complex systems, Critical infrastructure, Resilience

Introduction

As real-time data sources increase rapidly, the importance of detecting anomalies in streaming data grows significantly. Critical applications such as preventative maintenance (Kamat and Sugandhi 2020), fraud prevention(Pourhabibi et al. 2020), and fault detection (Jin et al. 2016) are found across diverse industries. Real-time anomaly detection presents unique challenges, nevertheless, the necessity for unsupervised, automated

systems that can continuously adapt to evolving data patterns is crucial, given the impracticality of human intervention.

The focus of the current study is energy systems (SES). Disruptions in SES, pose significant risks, including power outages (Toshev 2016), financial losses (Kamat and Sugandhi 2020), impacting sectors like healthcare (Stewart and Stewart 2024), Banking (Shibu et al. 2024). These disruptions stem from natural disasters, cyber-attacks, equipment failures, or human errors, leading to economic and environmental damage as well as compromised energy infrastructure service continuity. Early detection is vital for ensuring energy supply, and reliability, minimizing disruption impacts, maintaining efficiency, preventing safety hazards, and enhancing system resilience against cyber threats, thus securing the energy infrastructure (Vegesna 2024; Yao et al. 2024; Mazumder et al. 2024).

Data-driven anomaly detection in energy systems (ES) is essential as ES evolves into smart CPS Systems (CPSS) (Klaes et al. 2020). As Critical Infrastructures (CIs), energy systems play a vital role in maintaining societal functions. Therefore, it is crucial to detect any anomalies promptly to ensure the reliability and continuity of the vital services that ES provides to society (Narayan et al. 2023). This is the main motive for conducting the current systematic literature review to identify the trends, opportunities, and challenges to open future study lines in this domain.

This article contributes various aspects to advance the field of real-time anomaly detection in smart energy critical infrastructures. The study explores anomaly types that machine learning can detect and the factors influencing its success. the article examines deploying these algorithms at the Edge for efficiency and lower latency, alongside essential security measures. The discussion covers the role of advanced machine learning models, performance improvement methods, and the methodologies and tools used in the field. It also outlines areas for future studies to achieve accurate anomaly detection, addresses scalability, reliability, and service continuity challenges, and identifies enabling technologies for real-time implementation.

The paper is organized into five sections as follows: Sect. "A general background" covers general background information on real-time anomaly detection and provides preliminary definitions. Section "Systematic review methodology" offers a detailed description of the systematic review process, including the formulation of research questions, the selection of primary studies, the databases used, search queries, as well as the criteria for inclusion, exclusion, and quality assessment. Section "Quantitative and statistical analysis of selected papers" presents the search results and general statistics about the final list of the selected articles. Section "In-depth synthesis of results and discussion" synthesizes the review data, addressing the research questions. Finally, Sect. "Conclusion" concludes the paper and discusses the study's limitations and potential validity threats.

A general background

This section dives into the background knowledge about the complex world of NextGen energy infrastructures, where cyber (information technology), physical (hardware and grids), and social (human behavior and societal needs) aspects converge. Subsequently, the challenges that arise from this intricate interplay and delve into data-driven solutions for anomaly detection that can address issues within the context of Smart Energy Infrastructures are explored.

CPS challenges in nextgen energy infrastructures

In the context of NextGen energy infrastructures, Cyber-Physical Systems (CPS) integrate advanced technologies, physical energy infrastructure, and the social behaviors of energy consumers and producers, presenting unique challenges for sustainable operation. The dynamic nature of energy generation, distribution, and consumption necessitates continuous monitoring and adaptive control strategies to maintain stability and optimize performance (Sankey et al. 2014). For instance, in microgrids, real-time monitoring of power generation, consumption, and grid parameters is crucial for efficient and reliable operation (Veerakumar et al. 2023). Similarly, in large-scale power systems, real-time state estimation using phasor measurement units (PMUs) is essential for monitoring and controlling the grid (Mak-Hau et al. 2022). The complexity of the system even increases in the presence of prosumers and decentralized energy generation (Weinand et al. 2020). Due to this complexity, data-driven solutions offer promising approaches for ensuring service continuity. They excel in detecting anomalies and potential disruptions, enabling timely responses (Wang et al. 2024; Chou and Telaga 2014a; Capozzoli et al. 2018; Singh et al. 2024).

Another significant challenge is integrating diverse technologies and data sources. NextGen energy infrastructures involve a variety of technologies, such as renewable energy sources, energy storage systems, and smart grid devices Urishev (2019); Weinand et al. (2020); Schäfer et al. (2011); Leal-Arcas et al. (2020). Ensuring their seamless interoperability requires standardized protocols, communication interfaces, and data exchange formats (Mak-Hau et al. 2022). Additionally, the vast amounts of data generated by these technologies need to be effectively managed and analyzed to derive meaningful insights and enable informed decision-making. Considering the socio-ecologic and socio-technical evolution of energy infrastructures, data-driven solutions, and real-time anomaly detection can address the challenge of integrating diverse technologies and data sources by ensuring seamless interoperability through standardized protocols and efficient data management.

Finally, the social aspect of NextGen cyber-physical-social energy infrastructures also presents significant challenges. Human behavior and social interactions are critical in shaping energy consumption patterns and the adoption of new energy technologies. For example, engaging consumers in demand response programs and promoting energy-saving behaviors can significantly impact overall energy consumption and grid stability (Pan et al. 2022; al Rashid et al. 2022; Yin et al. 2022). Furthermore, regulatory and cybersecurity frameworks are paramount as infrastructures become more interconnected and reliant on digital technologies. The integration of distributed energy resources and the use of advanced metering systems introduce new vulnerabilities, while the implementation of machine learning and artificial intelligence in energy management systems raises concerns about data privacy and potential cyber-attacks. Balancing innovation with security is crucial. Therefore, studying adaptable and comprehensive regulatory frameworks is essential for fostering a secure and resilient energy ecosystem (Yin et al. 2022; Ekti et al. 2022). Addressing these cyber-physical-social challenges through real-time anomaly detection algorithms in a trustable environment requires a multidisciplinary approach to create a more resilient, efficient, and sustainable energy future. In the next subsection, the necessity of investigating real-time anomaly detection in SES will be detailed.

Anomaly detection for smart energy infrastructures

With the rapid advancement of information technology and the Internet of Things (IoT), traditional power grids are evolving into Smart Grids (SGs) (Xu et al. 2019). SGs feature Advanced Metering Infrastructure (AMI), which facilitates two-way communication between utility providers (UPs) and consumers, enabling functions like automatic meter reading and demand response (Yip et al. 2018a).

While traditional monitoring methods like checklists and even predictive maintenance fall short of providing real-time insights and enabling agile responses for smart grids (Colak et al. 2016), these very systems are fundamentally reliant on data-driven solutions for control. This underscores the critical role of real-time, data-driven anomaly detection for smart energy infrastructures.

By leveraging data-driven real-time anomaly detection, utilities can gain valuable insights from this vast amount of data. This enables them to identify and address issues like meter malfunctions (Yip et al. 2017), potential energy theft (Haq et al. 2023), and unexpected surges in demand (Tang et al. 2023). Early detection and response to these anomalies can improve grid reliability (Olatunde et al. 2024), prevent outages (Khediri and Laouar 2018), optimize energy consumption (Albogamy et al. 2022), and ultimately lead to a more resilient and secure smart energy infrastructure. In a nutshell, conducting the current systematic review on anomaly detection in smart energy infrastructures is fundamental to comprehensively understanding current methodologies, identifying gaps in existing research, and guiding the development of more robust and effective solutions for ensuring the resilience and efficiency of future SES.

Systematic review methodology

This section outlines the methodology used to conduct a systematic literature review (SLR) on the design, development, and implementation of real-time anomaly detection, specifically within the energy domain. The review aims to evaluate and interpret all available research relevant to the research questions outlined in sects. "Planning phase".

The research design of this study is Mixed-Methods Research Design including quantitative and qualitative synthesis of the results. Longitudinal Design features are considered to involve studying anomaly detection over an extended period to track changes or developments over time. Finally, the extracted articles (79 articles) are subjected to Descriptive, Exploratory, and Comparative study to focuses on describing real-time anomaly detection state of the art, challenges and opportunities. Figure 1 outlines the systematic plan for conducting the study, encompassing key elements such as research resources and objectives, methodology, sampling strategy, data collection methods, and analysis techniques. Finally, PRISMA Meta-data analysis for the systematic literature review is illustrated in the Fig. 2.

Planning phase

To establish the necessity for a systematic review, we examined existing literature on real-time anomaly detection from the SCOPUS digital library. The focus was on

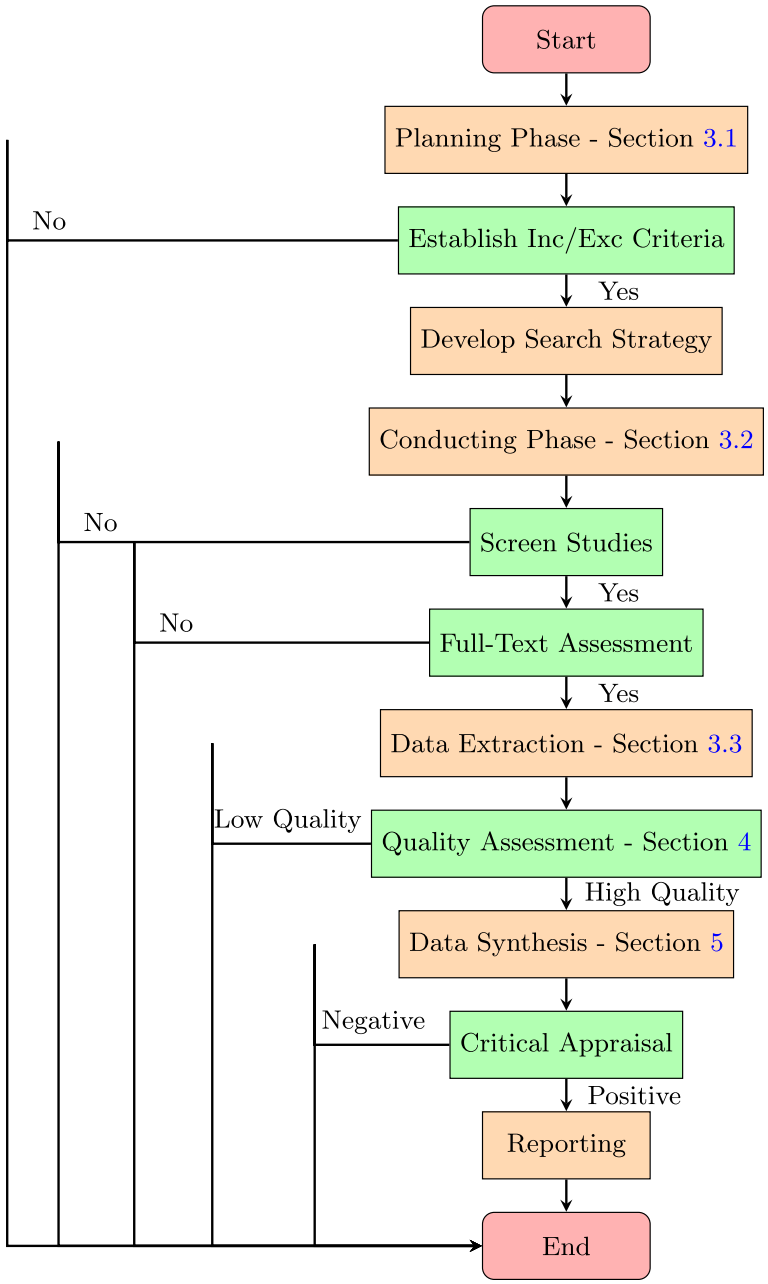

Fig. 1 Research design flowchart for systematic literature review

articles published in English, specifically in the subject area of Energy Systems. Our preliminary search revealed a gap in the literature concerning the development and deployment, and performance of real-time anomaly detection systems in energy applications.

The following research questions (RQs) were formulated to guide our review:

- *RQ1:* What types of anomalies can be detected using real-time machine learning algorithms, and what methodologies and tools are most effective for achieving accurate detection?

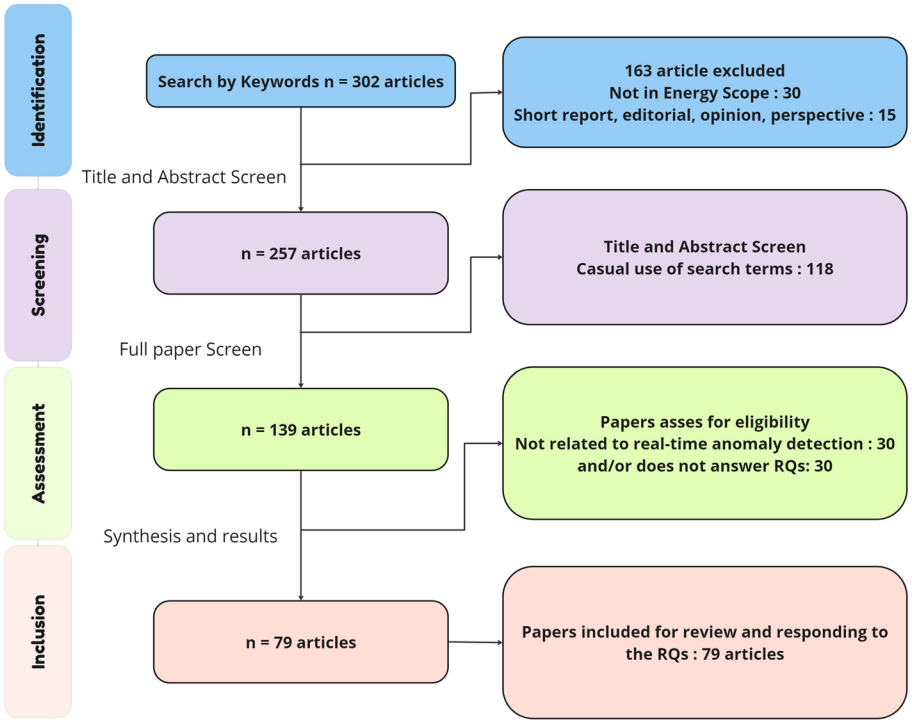

Fig. 2 Flow chart of the PRISMA Meta-data analysis for the systematic literature review

- *RQ2:* What key factors and methods contribute to the success or failure of real-time anomaly detection, and how can performance be optimized?
- *RQ3:* How can anomaly detection algorithms be effectively deployed across the Edge-cloud continuum, and what enabling technologies support these deployments?
- *RQ4:* What security measures are essential for protecting real-time anomaly detection systems from cyber threats?
- *RQ5:* What advancements and future directions are needed to enhance the accuracy and effectiveness of real-time anomaly detection systems?

A review protocol was defined to reduce the risk of researcher bias. The protocol includes criteria for article collection, selection, quality assessment checklists, data extraction, and synthesis strategies. This approach ensures the review is objective and transparent, enhancing the validity and reliability of the results.

Conducting phase

The conducting phase of this systematic review involved several steps: article collection, article selection, screening process, and quality assessment.

Article collection

Articles were collected from the SCOPUS digital library using predefined search criteria and keywords related to real-time anomaly detection in the energy domain. The search strategy involved decomposing research questions into individual terms,

Table 1 Article selection criteria

Criteria	Details
Primary exclusion	Limit subject area to Energy, language to English, exclude non-research articles, and casual use of terms.
Primary inclusion	Includes peer-reviewed documents, various document types, and a focus on real-time anomaly detection.
Stage 2 inclusion	Include documents that answer at least one research question.

Table 2 Screening stages

Stage	Description
Stage 1: Abstract Read	Initial screening based on abstract content.
Stage 2: Full Paper Study	In-depth review to ensure inclusion criteria are met.
Stage 3: Snowballing	Identify additional articles by reviewing references from selected papers.

collecting keywords from known primary articles, identifying synonyms, and defining search strings using Boolean operators.

Article selection

The selection and filtration process involved three stages, as detailed in Table 1.

Screening process

The screening process involved three stages, as detailed in Table 2.

Quality assessment

The quality of the selected articles was assessed using adapted questions rated on a scale from 0 to 1 (Yes = 1, Partially = 0.5, No = 0). Articles scoring at least 6 out of 10 were included. The assessment focused on ten criteria, as shown in Table 3.

Data extraction and synthesis

For each selected paper, data were extracted on anomaly types detected, contributing factors to detection success or failure, deployment strategies at the Edge, essential security measures, use of advanced machine learning models, performance enhancement methods, methodologies and tools used, future advancements and studies needed, scalability, reliability, and service continuity challenges, and enabling technologies for real-time anomaly detection. The extracted data were synthesized to answer each research question, providing a comprehensive overview of the current state of real-time anomaly detection in the energy domain.

Table 3 Quality assessment criteria

Criterion	Description
Clarity of research questions or objectives	Assess if the research questions or objectives are clearly stated.
Contextual placement within other studies	Evaluate how well the study is placed in the context of existing research.
Definition of detection or prediction methods	Determine if the methods used for detection or prediction are clearly defined.
Comparison of proposed methods with other techniques	Check if the proposed methods are compared with other existing techniques.
Specification of method parameters	Verify if the parameters of the methods are specified.
Availability of datasets	Confirm if the datasets used in the study are available.
Description of anomaly data sources	Examine if the sources of anomaly data are described.
Definition of features used	Assess if the features used in the study are defined.
Justification of evaluation metrics	Ensure that the evaluation metrics are justified.
Evidence-based answers to study questions or objectives	Determine if the study provides evidence-based answers to the research questions or objectives.

Supporting methods and tools

For text classification task (in this case topic classification), BERT transformer is used (Zhu et al. 2020). These algorithms learn patterns in text data to classify them into predefined categories (Van Aken et al. 2019). The next step of data analysis, text summarization, involves condensing a piece of text while retaining its key information. Generative AI is used for this part(Alli-Balogun 2024). Extractive summarization algorithms select and stitch together important sentences or phrases from the original text. Techniques include ranking sentences by importance (using graph-based methods that are very common in Large Language models (Nguyen and Nguyen 2015; Mihalcea and Radev 2011)) and then selecting top-ranked sentences. Abstractive summarization algorithms generate summaries that may contain new phrases and sentences not present in the original text.

Quantitative and statistical analysis of selected papers

In this section, we present a detailed quantitative analysis of the selected documents, emphasizing publication trends, key technologies, and application domains within the domain of real-time anomaly detection in energy systems. This comprehensive review aims to provide a clear understanding of the evolving landscape in this critical research area and identify key trends and potential gaps that warrant further investigation.

Figure 3 illustrates the annual trends in the number of papers published on real-time anomaly detection in energy systems from January 2014 to May 2024. The data reveals a notable upward trajectory, reflecting a growing scholarly interest in this field. The number of publications saw a modest increase from 2014 to 2017, followed by a significant surge starting in 2018, peaking in 2022. This sharp rise aligns with the increasing global emphasis on enhancing the resilience and efficiency of energy systems through advanced technological interventions. The slight decline observed in 2023 and early 2024 may be attributed to the lag in the publication process or a shift in research focus towards other emergent areas within the broader scope of energy systems.

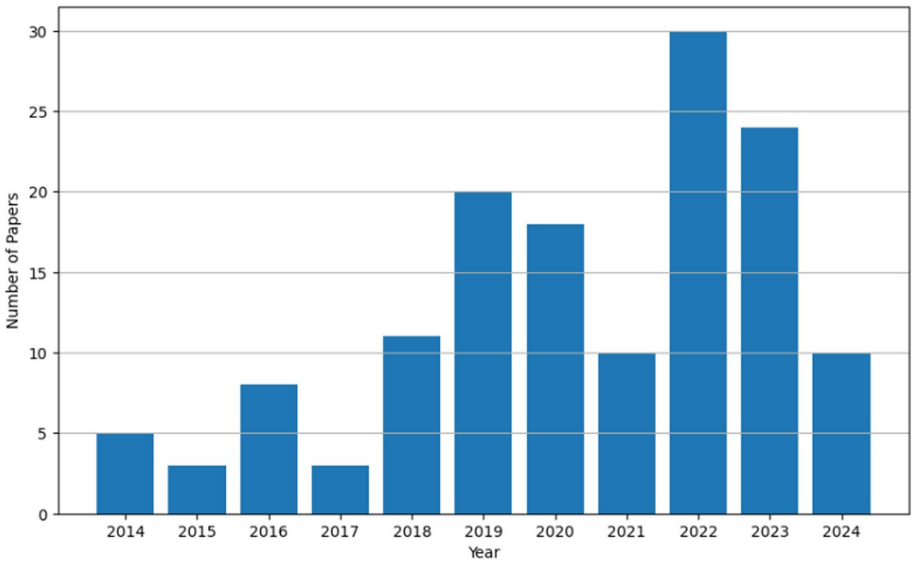

Fig. 3 Trend of papers published on real-time anomaly detection in energy systems by year (January 2014–May 2024)

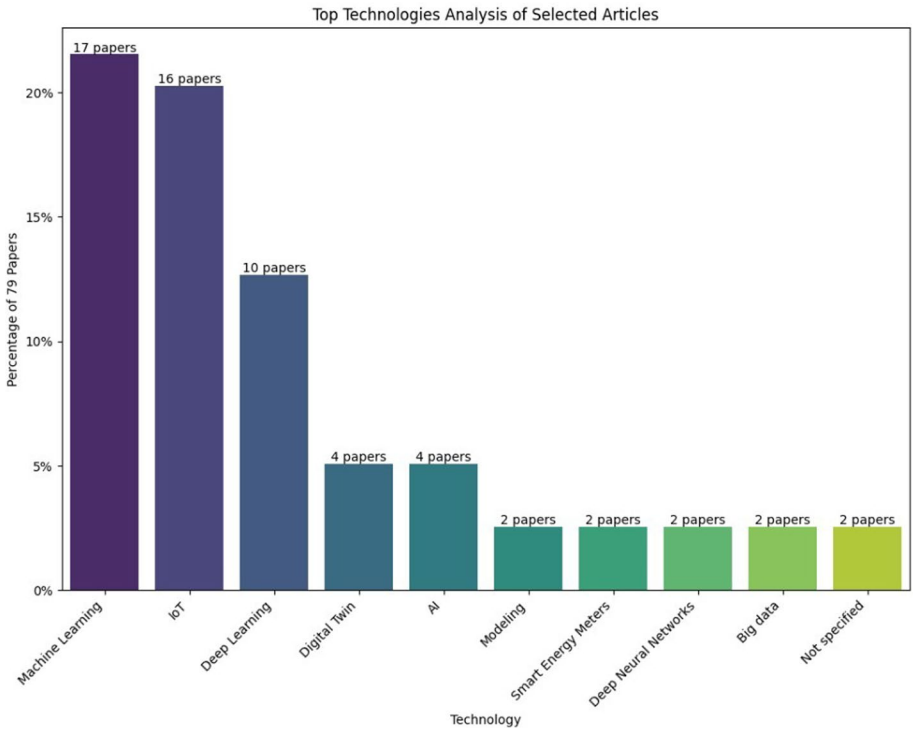

Fig. 4 Technologies used in the collected papers of real-time anomaly detection applications (top 10 highly cited technologies)

Figure 4 provides a statistical breakdown of the technologies utilized in the collected papers. The analysis identifies Machine Learning (ML) as the most dominant technology, being utilized in approximately 23.3% of the reviewed papers. This prominence is

likely due to ML's robust capability to analyze vast datasets and its adaptability in developing models that can detect complex patterns indicative of anomalies. The Internet of Things (IoT), featured in 21.9% of the papers, underscores the importance of real-time data collection and monitoring through interconnected devices, which is crucial for timely anomaly detection. Deep Learning (13.7%) also plays a significant role, reflecting its advanced capabilities in handling high-dimensional data and extracting intricate features for anomaly detection.

Other noteworthy technologies include Digital Twins (5.5%), Artificial Intelligence (AI) (5.5%), and Big Data Analytics (2.7%). These technologies highlight the integration of real-time simulation, predictive analytics, and large-scale data processing in modern anomaly detection systems. The diversity in the technological approaches, as depicted in Fig. 4, showcases the multidisciplinary nature of this research field and the varied methodologies employed to enhance the accuracy and efficiency of anomaly detection in energy systems.

The analysis of application domains, as shown in Fig. 5, reveals that a substantial portion of the research (30.9%) is concentrated on Power Generation. This focus highlights the critical need for ensuring operational reliability and safety in power generation facilities, where undetected anomalies can lead to severe disruptions and safety hazards. Power Distribution (21.3%) also emerges as a significant area of application, reflecting the ongoing efforts to improve the monitoring and control of distribution networks to prevent outages and maintain service quality.

The category labeled 'Others' (29.4%) includes diverse applications such as industrial automation, environmental monitoring, and smart city infrastructure, indicating a broad scope for anomaly detection technologies beyond traditional energy domains. Smaller

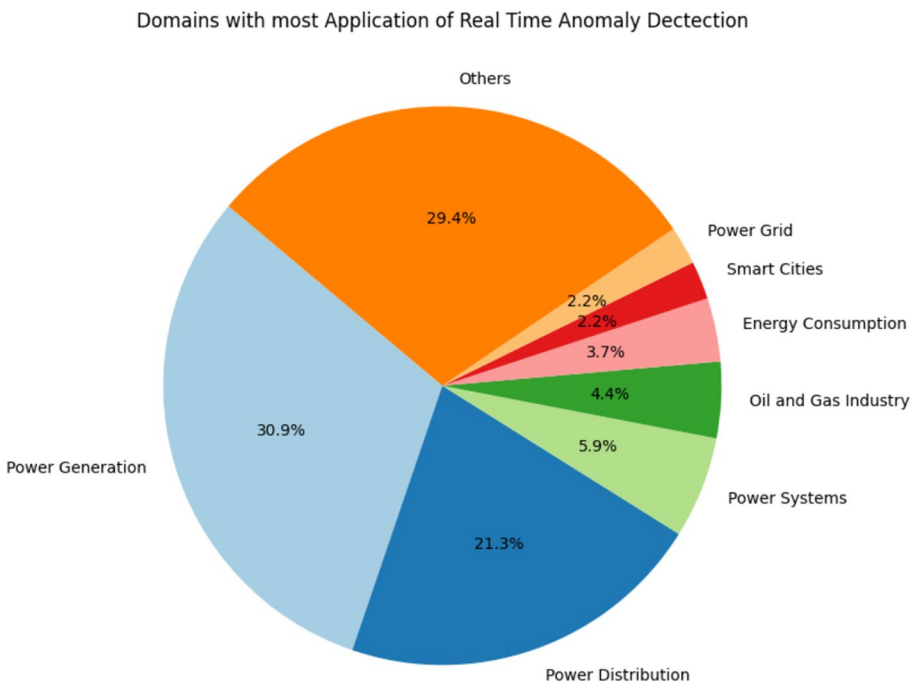

Fig. 5 Distribution of application domains in real-time anomaly detection

but noteworthy segments include Power Systems (5.9%), Oil and Gas Industry (4.4%), Energy Consumption (3.7%), Smart Cities (2.2%), and Power Grid (2.2%). These findings suggest that while the core focus remains on power generation and distribution, there is a growing interest in extending these technologies to other critical infrastructures, which are equally susceptible to anomalies that could impact operational efficiency and safety.

Overall, this statistical analysis provides a comprehensive overview of the current state of research in real-time anomaly detection for energy systems. It highlights the predominant technologies and application areas while also identifying gaps and opportunities for future research. The continued evolution of this field is likely to be driven by advancements in ML, IoT, and AI technologies, alongside increasing applications in diverse domains beyond traditional energy systems, as researchers seek to enhance the resilience and efficiency of critical infrastructures worldwide.

In-depth synthesis of results and discussion

Real-time anomaly detection

Table 4 presents a comprehensive overview of various anomalies within different categories of energy and power systems, along with the specific detection algorithms employed to identify these anomalies in the literature. The table aims to highlight the diverse types of anomalies that were subject to research. It also underscores the various methodological approaches taken by researchers to detect and address these anomalies, showcasing the extensive efforts in enhancing the reliability, security, and efficiency of energy systems.

The analysis of the Table 4 reveals a distinct trend in the research focus on particular types of anomalies within the energy and power systems domain. This trend highlights the specific anomalies that are most frequently studied by researchers. Among the identified categories, power system faults—including power supply issues, faults in grid voltage, and short, duration voltage anomalies, etc—emerge as the most highly studied anomalies (Fig. 6 shows the highly studied anomalies).

This prevalent focus suggests a significant research interest in ensuring the reliability and stability of power systems, likely due to their critical impact on energy distribution and consumption. Additionally, other anomalies such as electricity theft, cybersecurity threats, and faults in solar panels and batteries also receive considerable attention in Fig. 6. This pattern underscores the researchers' tendency to prioritize anomalies that directly affect the efficiency, security, and sustainability of energy systems, reflecting broader industry and societal concerns.

In order to provide a deeper understanding of the practical applications of ML in realtime anomaly detection (classified in Table 4), a detailed case studies with reusable approach from each anomaly category has been selected for further exploration. Below, the applications of these algorithms are briefly described to highlight their significance in different contexts.

Table 4 Categorization of anomalies in energy and power systems with corresponding detection algorithms

Anomaly category	Ref.	Exact type of anomaly	Detection algorithm
Electric power generation and distribution	(Jiang et al. 2022a)	Electricity theft detection	Isolated forest algorithm
	(Bhaskar et al. 2023)	Power theft or faulty meters, and sensor or communication failures	N/A
	(Mallor et al. 2017)	Detecting outliers or faults in the solar energy production of identical sets (sister arrays) of photovoltaic (PV) solar panels	N/A
	(Shapsough et al. 2020)	Shading anomalies on solar panels	Isolation Forest algorithm, Siamese neural network, latent space, NN and kNN
	(Leng and Qiu 2023)	Anomalies in power system operation data, including parameter errors, bad measurement data, and topology errors	Fuzzy C-means algorithm
	(al Rashid et al. 2022)	Detecting false data injection attacks in smart grid systems	CNN-LSTM based auto-encoder
	(al Rashid et al. 2022)	False data injection attacks in the context of the research	CNN-LSTM based auto-encoder
	(Rosch et al. 2019)	Load drop attacks, transient events in voltage signals, harmonic distortions, asymmetries in three-phase systems, and deviations in voltage gradients within distribution grids	N/A
	(Pei et al. 2022)	Anomalies in power grid data	Improved Support Vector Machine (SVM) and random forest
		Short-duration voltage anomalies in smart grids	N/A
	(Yen et al. 2019)	Anomalies in power generation measurements	N/A
	(Mak-Hau et al. 2022)	Deviations from baseline bidding behaviors in smart grids indicating cyber threats	N/A
	(Nur et al. 2019)	Anomalies in weighted graphs and multidimensional time series	DynWatch, a domain knowledge-based and topology-aware algorithm for anomaly detection using sensors placed on a dynamic grid
	(Li et al. 2022)	Anomalies in the power supply system	Seasonal Autoregressive Moving Average (SARMA)
	(Abedi et al. 2023)	Energy theft and meter irregularities in AMI and smart grids	LR

Table 4 (continued)

Anomaly category	Ref.	Exact type of anomaly	Detection algorithm
	(Yip et al. 2018b)		N/A
		Reactor power operation deviations	Hidden Markov Modeling (HMM)
	(Ayaz et al. 2003)	Thermoacoustic instabilities in combustors	Hybrid anomaly detection model involving multivariate linear regression, response surface methodology, and multi-layer perceptron
	(Mondal et al. 2019)	Anomalies in high-power generator vibrations	Light weight algorithms
	(Kirbaş and Kerem 2021)	Detection of disturbances and transients in smart grids	N/A
	(Ekti et al. 2022)		N/A
Road lighting and distribution systems	(Śmiałkowski and Czyżewski 2022)	Anomalies in road lighting systems, including lamp failures, schedule deviations, and energy theft	One based on an autoregressive integrating moving average periodic model (SARIMA) and the other based on a recurrent network (RNN) using long short-term memory (LSTM)
Battery faults and performance	(Jiang et al. 2022a)	Battery voltage faults in electric vehicles	Isolated forest algorithm
	(Bhaskar et al. 2023)	Internal short circuits (ISC), air-flow anomalies, loose temperature and voltage sense leads, and voltage dropouts in battery systems	N/A
	(Jiang et al. 2022b)	Short circuit faults, open circuit faults, and other early battery faults in lithium-ion batteries	N/A
	(Cadini et al. 2019)	Drops in battery capacity, unexpected changes in degradation dynamics, abnormal end-of-life predictions, deviations from normal operating conditions, and any irregularities impacting battery performance and safety	Combination of the use of neural networks (MLP, Multi-Layer Perceptron) ad particle filters (PF)
	(Noureen et al. 2019)	Faults in grid voltage and frequency phasors	N/A
	(Pileggi et al. 2019)	The running mode, and the operating mode	N/A
	(Ayaz et al. 2003)	Anomalies related to operational deviations in reactor power operation	N/A
Abnormal pattern detection	(Cheng et al. 2023)	Pipe bursts and situations causing abnormal sensor data in distribution systems	Local Outlier Factor (LOF) algorithm

Table 4 (continued)

Anomaly category	Ref.	Exact type of anomaly	Detection algorithm
	(Wong et al. 2021)	Real-time machine learning algorithms detecting abnormal water quality parameters like turbidity and water level	N/A
Resource extraction and transportation	(Ramba et al. 2021)	Real-time machine learning algorithms detecting drill string failures during rotary drilling operations	Mathematical Model
	(Giunta et al. 2019)	Anomalies in fluid transportation pipelines (leaks, blockages, equipment failures)	N/A
Industrial production and maintenance	(Guillen et al. 2020)	Analyze the failures of two drywell cooling fans at a nuclear power plant	Long Short-Term Memory (LSTM) recurrent neural network, used in conjunction with the RELAP5-3D model
	(Chen et al. 2022b)	Anomalies in industrial fields (intrusion detection, video anomaly diagnosis, polluted gas recognition, gas production monitoring)	Neural network with adaptive threshold
	(Chakraborty et al. 2008)	Detecting damage in refractory walls of gasification systems, using temperature profiles and dynamic responses	N/A but is a damage prediction algorithm based on an integrated simulation model and the utilization of a pattern identification tool called "symbolic dynamic filtering" (SDF)
	(Yin et al. 2023)	The real-time detection of the state of connecting bolts of the rotor coil in a hydro generator unit	Local enhancement and regional features
	(Jiang and Zhao 2022)	micro-cracks in PV module cells, surface defects in industrial products, pavement surface anomalies, and concrete bridge cracks	attention classification-and-segmentation network
	(Ramesh et al. 2022)	detection of abnormalities in solar power plants	N/A
	(Ramesh et al. 2022)	condition changes in transformers	N/A
Monitoring and security	(Pandit and Infield 2018)	real-time machine learning algorithms detecting false data injection attacks in Plug-in Electric Vehicles (PEVs) integration within smart grids	N/A
	(Ramesh et al. 2022)		N/A

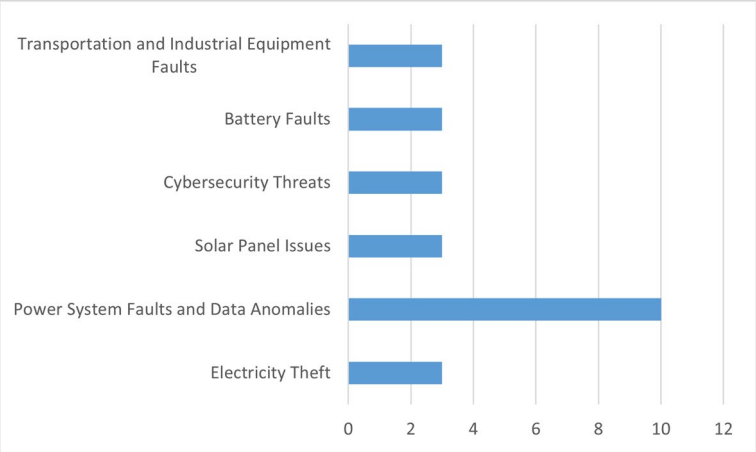

Fig. 6 Distinct trend in the research focus on particular types of anomalies inferred from Table 4

- *Electric Power Generation and Distribution:*

  Reference: Jiang et al. (2022a)

  Application: Electricity Theft Detection

  This paper introduces a method for diagnosing faults in power lithium batteries using an isolation forest algorithm. The process begins by applying Variational Mode Decomposition (VMD) to the voltage data collected by the Battery Management System (BMS). This step extracts static components linked to aging inconsistencies and dynamic components that indicate abnormal behavior. Next, key features such as autocorrelation and cross-correlation are identified and fed into the isolation forest algorithm, which then detects faulty battery cells. Figure 7 illustrates the complete workflow of this fault diagnosis approach using the isolation forest algorithm. The approach of this case-study can be used for utility companies as it helps identify irregular consumption patterns and fraudulent activities, thereby preventing revenue loss and improving the efficiency of energy distribution.
- *Battery Faults and Performance:*
  Reference: Cadini et al. (2019)
  Application: Battery Performance Monitoring
  This research focuses on detecting irregularities in battery performance and safety using a combination of neural networks and particle filters. Equation 1 shows the diagnostic index based on the particle mean likelihoods, which is introduced by Cadini et al. (2019) and can be implemented in electric vehicles and energy storage systems, ensuring reliable battery operations and extending the lifespan of these systems:

$$\mathrm{LLR}_k = \ln\left(\frac{\frac{1}{N_s}\sum_{i=1}^{N_s}\mathscr{L}_{k-1}^{(i)}}{\frac{1}{N_s}\sum_{i=1}^{N_s}\mathscr{L}_{k}^{(i)}}\right) \tag{1}$$

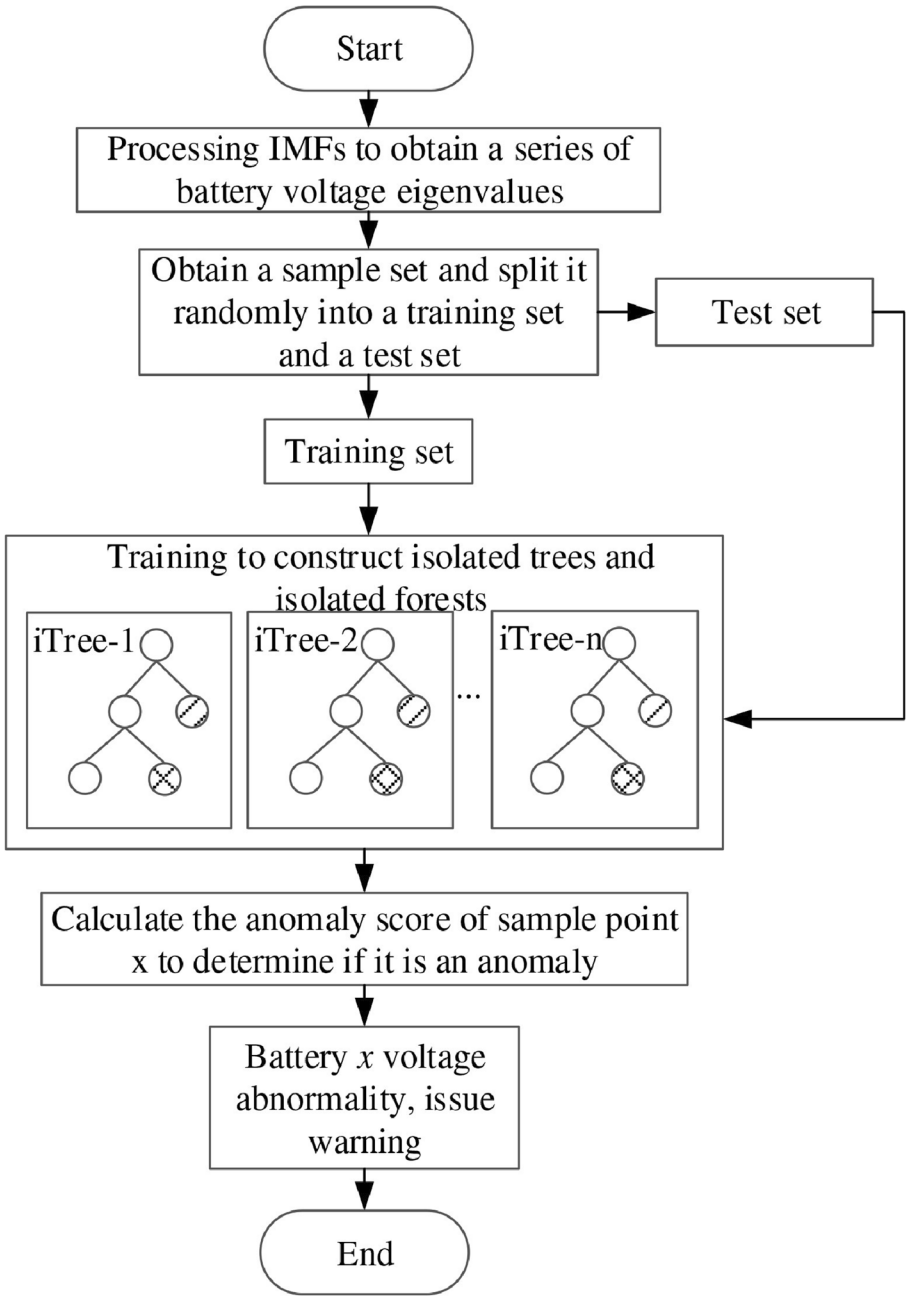

Fig. 7 Workflow of this fault diagnosis approach using the isolation forest algorithm proposed by Jiang et al. (2022a)

where the numerator and the denominator are the mean likelihoods of the particles (i.e., the MLP network parameters) at the cycles $k-1$ and k, respectively. Operatively, if the mean likelihood of the particles at cycle k is equal to that at cycle $k-1$, i.e., $\mathscr{L}_k = \mathscr{L}_{k-1}$, then $LLR_k = 0$ and, probably, the observed capacity behavior does not differ significantly from that predicted by the N_s particles, i.e. the N_s MLP networks. At each cycle k the posterior pdf of the end of life time, $p(\mathrm{EOL_k} \mid \boldsymbol{z}_{1:\mathrm{k}})$, is estimated by letting the N_s MLP networks, associated to the N_s parameter samples

(particles) $\boldsymbol{x}_k^{(i)}, i = 1, \cdots, N_s$, evolve until their capacity predictions reach the predefined threshold, following a simple particle projection strategy.

- *Abnormal Pattern detection:*
  Reference: Cheng et al. (2023)
  Application: Anomaly Detection in Distribution Systems
  The study uses the Local Outlier Factor (LOF) algorithm to detect abnormal sensor readings in the distribution networks. This paper uses the K-means clustering algorithm to propose an improved algorithm K-LOF of the density-based local abnormal factor detection algorithm LOF, and optimizes the neighborhood query process. Figure 8 shows flow chart of abnormal information detection; and Fig. 9 illustarates Large-scale network abnormal behavior detection based on cluster pattern recognition that can be used for reducing the time complexity of the anomaly detection process while enhancing detection accuracy.
- *Resource Extraction and Transportation:*
  Reference: Ramba et al. (2021)
  Application: Anomaly Detection in Drilling Operations
  The paper employs mathematical models for real-time detection of drill string failures during rotary drilling operations. This application is essential in the oil and gas industry for preventing costly equipment damage and improving safety during resource extraction. The Decision Support System proposed by Ramba et al. (2021) is shown in Fig. 10. The proposed hookload model can be reused in detecting the downhole complications in real-time drilling operations to reduce the NPT.
- *Industrial Production and Maintenance:*
  Reference: Guillen et al. (2020)
  Application: Failure Analysis in Nuclear Power Plants
  This research utilizes Long Short-Term Memory (LSTM) recurrent neural networks in combination with the RELAP5-3D model to analyze cooling fan failures at a nuclear power plant. The application ensures the early detection of critical faults, enhancing plant safety and reliability. Guillen et al. (2020) proposed a methodol-

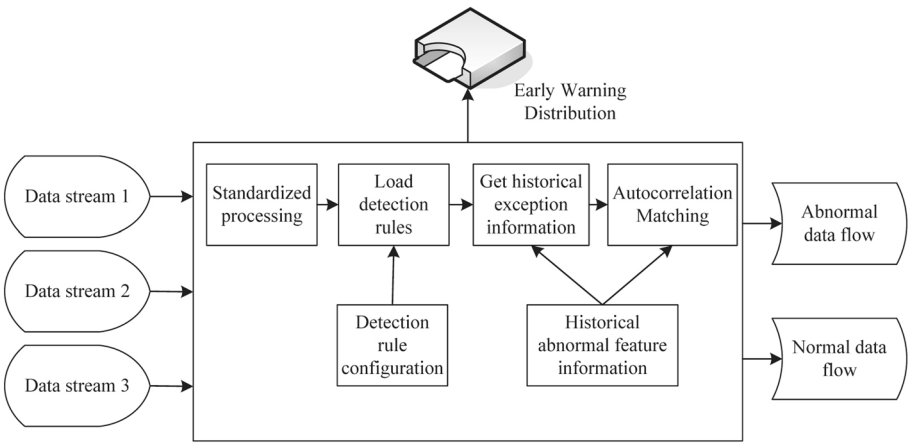

Fig. 8 Flow chart of abnormal information detection Cheng et al. (2023)

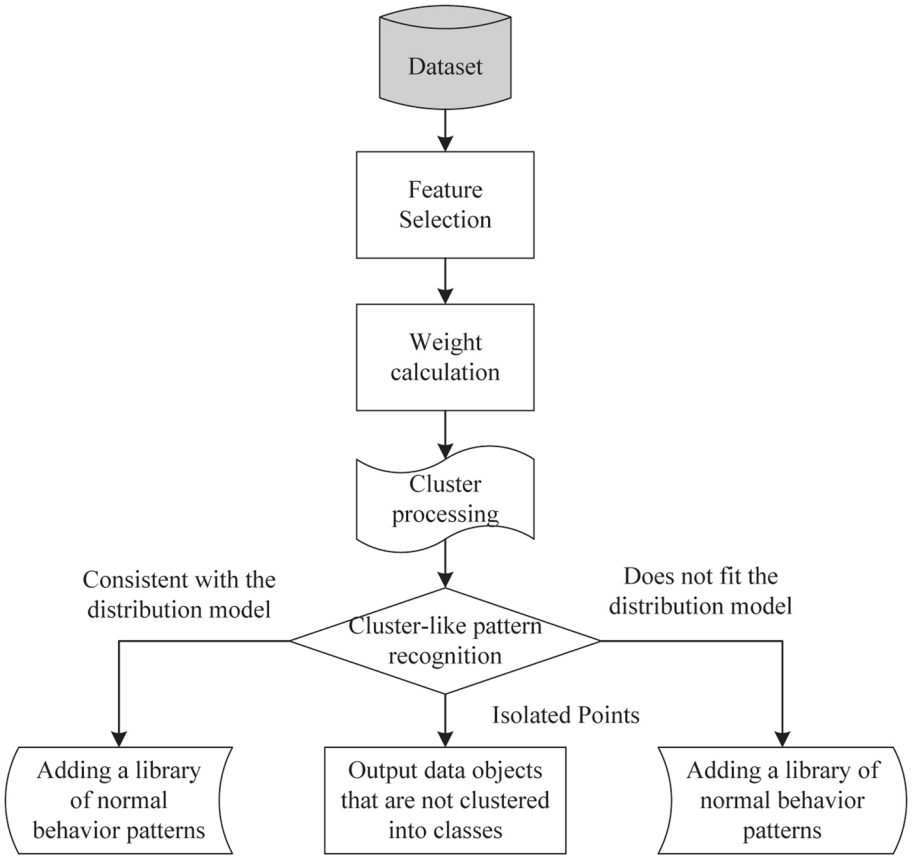

Fig. 9 Large-scale network abnormal behavior detection based on cluster pattern recognition (Cheng et al. 2023)

ogy named "RELAP5-3D solution methodology" (See Fig. 11) that showcased viable approach to predict the expected FCU outlet temperature.

This combined approach offers a valuable method for detecting anomalies using data from existing plant sensors. The main objective is to equip nuclear power plants with the ability to identify hidden issues within process data, anticipate potential catastrophic failures, and take preventive measures. By proactively identifying unusual system behaviors that might signal upcoming problems, operators can respond before serious issues arise. Approaches like this are essential for ensuring the long-term reliability and safety of light water reactors.

Factors in anomaly detection success

Studying the success factors of digital twins, even in its early stages, is crucial. Understanding what makes this technology thrive allows ongoing researches to avoid pitfalls and maximize the technology's potential. By analyzing past implementations, we can identify key areas that are essential for achieving the desired outcomes. This knowledge helps us develop a road-map for research society for successful integration, ensuring digital twins. Table 5 generalizes the success factors

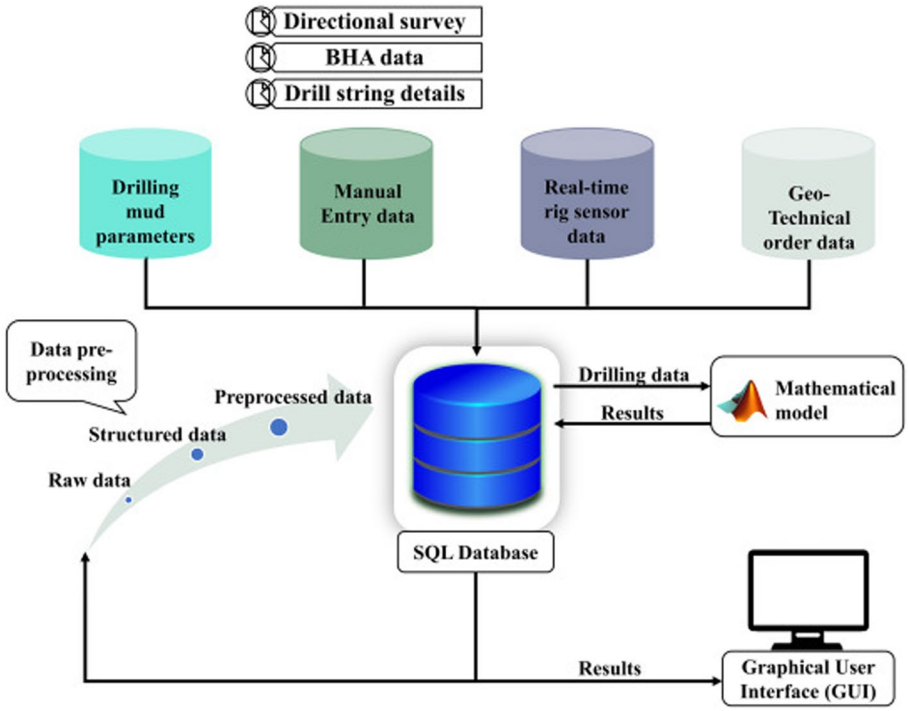

Fig. 10 Proposed decision support system for etecting the downhole complications in real-time drilling operations (Ramba et al. 2021)

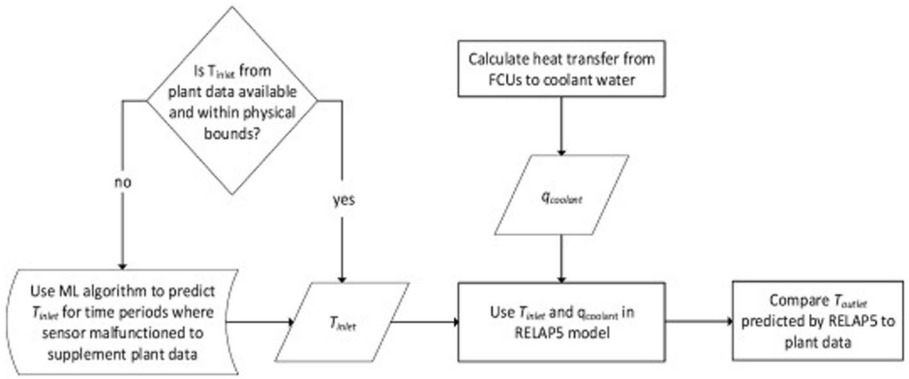

Fig. 11 RELAP5-3D solution methodology (Guillen et al. 2020)

that are mentioned in the articles into common terms (detailed in Secti.5.1) and lists the references that mention each factor.

Distributed deployment

Deploying anomaly detection systems benefits greatly from a hybrid edge-cloud approach. Edge computing allows for real-time data processing and immediate detection of anomalies close to the data source, minimizing latency and reducing bandwidth usage. Cloud computing, on the other hand, offers the scalability and computational power needed for analyzing large datasets and improving detection algorithms. Combining

Table 5 The success factors for real-time anomaly detection

Qty	References	Success factor	Description
12	(Hu et al. 2022; Śmiałkowski and Czyżewski 2022; Mohammadpourfard et al. 2021; Pandey et al. 2020; Kirbaş and Kerem 2021; Wang et al. 2023b; Pei et al. 2022; al Rashid et al. 2022; Singh et al. 2024; Xu et al. 2023; Guillen et al. 2020; Krishna et al. 2013)	Data Quality	Clean, accurate data feeds the digital twin, ensuring a reliable representation of the real system.
3	(Sleiti et al. 2022; Mak-Hau et al. 2022; Wyss et al. 2023)	Algorithm Choice and Efficiency	Choosing the right algorithms optimizes performance and allows the digital twin to react effectively to changes.
5	(Tehrani et al. 2022; Yin et al. 2022; Pan et al. 2022; Brahma et al. 2016; Vikram et al. 2020)	Computational Resources	Sufficient processing power ensures the digital twin can handle complex simulations and real-time data analysis.
6	(Yan et al. 2022; Sun et al. 2022; Pei et al. 2022; Davarifar et al. 2014; Yip et al. 2018b; Jiang and Zhao 2022)	Accuracy and Sensitivity	A high-fidelity digital twin minimizes discrepancies between the virtual model and the physical system.
6	(Wang et al. 2022; Wong et al. 2021; Pan et al. 2022; Vikram et al. 2020; Azhar et al. 2022; Xu et al. 2021)	Sensor Quality and Integration	Reliable sensors provide valuable data for the digital twin, and seamless integration ensures smooth information flow.
5	(Wen et al. 2022; Liu and Aldrich 2023; Wen et al. 2022; Jin et al. 2011; Sun et al. 2022)	Integration of Multiple Methods	Combining various data sources and analytical techniques creates a richer and more comprehensive digital twin.
5	(Baker and Shadmand 2023; Ramesh et al. 2022; Sun et al. 2022; Abedi et al. 2023; Sleiti et al. 2022)	Real-time Processing	Real-time data analysis enables the digital twin to reflect the physical system's current state and predict future behavior.
5	(Wyss et al. 2023; Jiang et al. 2022a; Wang et al. 2022; Mak-Hau et al. 2022; Tehrani et al. 2022)	System Robustness	A robust digital twin can withstand data fluctuations and unexpected events, and guarantee the resilience of DT.
6	(Abdelmoula et al. 2023; Wang et al. 2023a; Ramesh et al. 2022; Jiang and Zhao 2022; Wang et al. 2019; Ji et al. 2021)	Scalability	The ability to handle growing data volumes and evolving needs ensures the digital twin remains valuable in the long term.

both methods leverages the strengths of each, ensuring efficient data processing and robust anomaly detection capabilities.

Table 6 compares various studies on KPIs and deployment models for anomaly detection in edge-cloud environments.

The table reveals diverse KPIs such as accuracy, energy efficiency, and latency, reflecting the multifaceted evaluation of anomaly detection systems. It shows the frequent use of edge computing for real-time processing to address privacy and latency issues, and highlights hybrid models that integrate edge and cloud resources for improved system performance. These insights inform the trade-offs and complementarities between edge and cloud computing, guiding future system designs and research.

Based on KPIs that are used to evaluate the efficiency of the system, the highest used KPIs are: Cost Efficiency, resource Utilization, Operational Efficiency.

Table 6 Comparison of key performance indicators (KPIs) and deployment models for edge-cloud continuum in various research articles (part 1)

Article	KPIs	Edge-cloud deployment
(Alsalemi et al. 2022)	Accuracy of Appliance Identification, Accuracy of Anomaly Detection, Energy Savings, Data Collection Efficiency, Response Time, Cost Efficiency, User Engagement	Edge computing is used for processing tasks closer to end-users, enhancing real-time responses and reducing latency.
(Śmiałkowski and Czyżewski 2022)	Efficiency of energy, System reliability, Timeliness of intervention, Safety	Algorithms allow for practical industrial implementation with relatively low equipment requirements.
(Ekti et al. 2022)	Efficiency and Productivity, Quality Control, Resource Utilization	Digital twins are enhanced through a hybrid deployment integrating edge computing and cloud resources for real-time processing and scalability.
(Shoman and Burr 2023)	Accuracy and Uncertainty, Detection Sensitivity, Cost Efficiency	Utilizes unattended monitoring methods at the edge integrated with centralized cloud analysis for improved safeguards.
(Ramesh et al. 2022)	Fault Detection Time, Cost Savings	Utilizes a distributed architecture with IoT gateways and sensor modules for real-time data collection and processing.
(Wang and Bu 2020)	Latency, Throughput, Resource Utilization, Energy Efficiency, Quality of Service (QoS)	The design focuses on optimizing resource utilization across edge and cloud continuum for efficient service delivery.
(Bushehri et al. 2021)	Energy Consumption, Operational Efficiency, Over-Consumption Detection, Accuracy Metrics	Introduces a new architecture for real-time execution trace analysis using edge computing for immediate anomaly detection.
(Nur et al. 2019)	Detection Accuracy of Bad Data, Baseline Establishment, Real-Time Detection Capability, Impact on Operational Efficiency	Deploys edge computing for real-time data processing to detect and respond to anomalies, complemented by cloud-based analytics for historical data.
(Anagnostou et al. 2018)	Anomaly Detection Accuracy, Guaranteed Absence of False Alarms, Real-Time Implementation Suitability, System Model Adaptability	Uses edge computing for real-time monitoring and anomaly detection in dynamic operational environments, supported by cloud resources for data management and analytics.
(Giunta et al. 2019)	Anomaly Detection Accuracy, Sensor Data Integrity, Operational Efficiency Improvements, Integration of Data Sources	Implements edge computing for real-time monitoring along with cloud-based analytics for large-scale data management and predictive modeling.
(Krishna et al. 2013)	Occupancy Estimation Accuracy, Energy Wastage Tracking, Consumption Anomaly Detection, Visualization of Energy Breakdown	Employs edge technologies for real-time energy usage analysis and optimization, enhancing responsiveness and efficiency in energy management.
(Abedi et al. 2015)	Real-Time Monitoring KPI, Cybersecurity KPI	Integrates edge and cloud processing to manage complexity and scale in energy consumption models.

Since these 3 KPIs are identified as high-frequency KPIs across the analyzed articles, their prominence indicates their critical importance in evaluating and optimizing various systems and technologies discussed in the literature.

Cost Efficiency: Edge computing can reduce costs associated with data transmission and storage by processing data locally (Wang and Bu 2020), thus minimizing reliance

on expensive cloud resources (Nur et al. 2019). However, Alsalemi et al. (2022); Shoman and Burr (2023) raised the importance of system's cost-effectiveness while maintaining measurement accuracy. on the other hand, the effectiveness of the system in reducing energy consumption through the implementation of edge-cloud architecture in real-time monitoring is studied by Alsalemi et al. (2022).

Resource Utilization: Efficient use of edge and cloud resources ensures optimal performance and scalability of applications(Wang and Bu 2020; Giunta et al. 2019). It involves balancing workload distribution, minimizing latency, and maximizing the use of available computing power and storage capacity.

Operational Efficiency: edge-cloud continuum aims to enhance operational efficiency by leveraging edge computing for real-time data processing and decision-making (Bushehri et al. 2021; Nur et al. 2019; Giunta et al. 2019).

An in-depth investigation of the reported results shows significant improvement in various KPIs through edge-cloud deployment. However, while some articles provide quantitative measurements of these KPIs, others report improvements qualitatively. For instance, (Śmiałkowski and Czyżewski 2022) deployed the M2SP-EdgeIoE framework, which enables edge computing capabilities for applications such as data collection, energy disaggregation, and appliance identification. This system integrates mobile applications for automation and visualization, providing real-time responses to anomaly detection, achieving a 95% accuracy rate in detecting anomalies in energy consumption and a 98.49% accuracy in identifying appliances based on their energy consumption patterns.

Conversely, other articles report improvements qualitatively. For example, Śmiałkowski and Czyżewski (2022) describes enhancements in:

1. *Energy Efficiency*: Measured by the reduction in overall energy consumption of the lighting system.
2. *System Reliability*: Assessed by the ability to detect and resolve anomalies in the operation of street lamps.
3. *Safety*: Evaluated by the ability to ensure adequate and constant illumination of streets for the safety of pedestrians and vehicles.

Security of anomaly detection ecosystem

Anomaly detection systems are a critical shield against cyber-attacks for future digital infrastructures (Aghazadeh Ardebili et al. 2024). Unlike traditional methods that focus on known threats, anomaly detection identifies unusual activity in real-time, catching novel attacks and reducing false alarms. However, deploying them in a trusted environment is crucial (Eichler et al. 2024). This is especially important for Digital Twins and other future technologies that rely on anomaly detection as a microservice to safeguard the integrity of data and simulations.

On the otherhand, Data security, confidentiality, and integrity provide the detection ecosystem with accurate and complete information to identify the disruptions. In cyber security domain, for instace, the system might be working with misleading

Table 7 Anomaly detection security measures

Security measure	References
Data encryption	(Hu et al. 2022; Śmiałkowski and Czyżewski 2022; Chen et al. 2023, 2022b; Leng and Qiu 2023; Wong et al. 2021; Ramba et al. 2021; Guillen et al. 2020; Li et al. 2020; Kirbaş and Kerem 2021; Wang and Bu 2020; Xu et al. 2021; Yip et al. 2018b; Noureen et al. 2019; Funde et al. 2019)
Access control	(Hu et al. 2022; Śmiałkowski and Czyżewski 2022; Tehrani et al. 2022; Ramesh et al. 2022; Leng and Qiu 2023; Wong et al. 2021; Ramba et al. 2021; Guillen et al. 2020; Li et al. 2020; Kirbaş and Kerem 2021; Wang and Bu 2020; Xu et al. 2021; Yip et al. 2018b; Noureen et al. 2019; Funde et al. 2019)
Resource assessment	(Śmiałkowski and Czyżewski 2022)
Data confidentiality	(Wang et al. 2022; Sawas and Farag 2023; al Rashid et al. 2022; Guillen et al. 2020; Xu et al. 2021; Abedi et al. 2015)
Intrusion detection	(Hu et al. 2022; Ramesh et al. 2022; Wang et al. 2022; Veerakumar et al. 2023; al Rashid et al. 2022; Liu and Aldrich 2023; Wen et al. 2022; Ramba et al. 2021; Guillen et al. 2020; Li et al. 2020; Kirbaş and Kerem 2021; Wang and Bu 2020; Chahla et al. 2020; Yip et al. 2018b; Rosch et al. 2019; Hong et al. 2011)
Secure data collection	(Sawas and Farag 2023; Ramesh et al. 2022; Xu et al. 2023; Yip et al. 2018b; Funde et al. 2019; Shapsough et al. 2020)
Deployment in trusted environment	(Sawas and Farag 2023; Abedi et al. 2023; Chen et al. 2023; Xu et al. 2023; Leng and Qiu 2023; Wong et al. 2021; Pandey et al. 2020; Shapsough et al. 2020; Anagnostou et al. 2018; Rosch et al. 2019)
Secure communication	(Chen et al. 2023, 2022b; Ramesh et al. 2022; Xu et al. 2023; Wen et al. 2022; Śmiałkowski and Czyżewski 2022; Kirbaş and Kerem 2021; Xu et al. 2021; Chahla et al. 2020; Noureen et al. 2019; Rosch et al. 2019)
Compliance with standards	(Chen et al. 2022b; Ramesh et al. 2022; Guillen et al. 2020; Yip et al. 2018b; Wen et al. 2022; Abedi et al. 2015; Giunta et al. 2019)

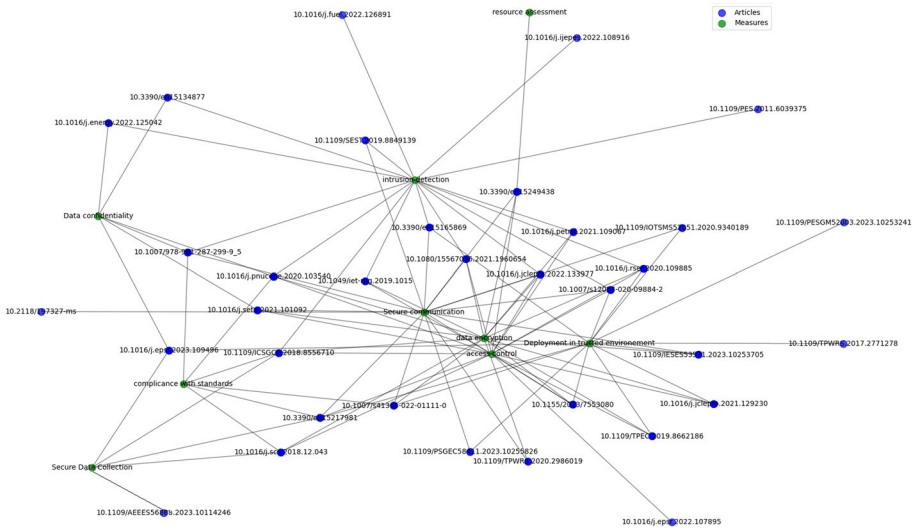

Fig. 12 Network graph of articles and security measures. Each node in the graph represents either a security measure or an article (identified by its DOI). The edges between the nodes indicate which articles employ which security measures. (The article nodes in blue are named by the DOI of the article in each node and the security measures are taken from Table 7)

clues or even fabricated evidence, making it impossible to detect a cyber attack. By protecting these critical entities from cyber threats, anomaly detection ensures the smooth operation of future digital infrastructures, acting as a vital service for society.

Table 7 lists the anomaly detection Security Measures to safeguard the anomaly detection ecosystem. These measures are extracted from the full article read of the final list.

The network graph of articles and security measures in Fig. 12 provides valuable insights into the interconnections between different security measures employed in the anomaly detection ecosystem. The insights inferred from the graph is detailed in the following categories: Central Measures, Article Overlap, Sparse Connections, Clustered Measures, Research Gaps.

Central measures: Security measures that have many connections (edges) to various articles are central in the graph. These central measures are the most commonly employed security measures across the studied articles. For instance, if "data encryption" and "access control" have numerous connections, it suggests that these measures are fundamental components in securing anomaly detection systems.

Article overlap: Articles connected to multiple security measures indicate a comprehensive approach to security. These articles implement a variety of security measures, suggesting a robust and multi-layered security strategy. Identifying these articles can highlight best practices and exemplary research in the field. Based on the network graph, here are five articles that exemplify best practice approach: (Hu et al. 2022; Ramesh et al. 2022; Sawas and Farag 2023; Chen et al. 2022b, 2023)

Sparse connections: Security measures with fewer connections may indicate niche or specialized measures that are less frequently employed but could be significant in specific contexts or emerging areas of research. The sparsely connected measures are: Secure Data Collection, Compliance with Standards, Resource Assessment. Future researchers can focus on these less-explored measures to contribute to the field's growth.

Clustered measures: Groups of security measures that are frequently implemented together can be identified through clusters in the graph. These clusters suggest common security strategies or frameworks that are popular in the literature. For example, if "intrusion detection," "secure communication," and "deployment in a trusted environment" frequently appear together, it indicates a common security strategy.

Research gaps: Measures with few or no connections might point to areas that are under-researched or newly emerging, suggesting opportunities for future work. Researchers can focus on these less-explored measures to contribute to the field's growth. The most significant gap is the implantation of Resource Assessment. Notwithstanding, assessing and managing data resources is essential in securing any data-driven systems (Goodhue et al. 1988; Graber and Kleinhammer 2016; Mead 1982).

In summary, the network graph not only highlights the most commonly used security measures. Notwithstanding, measures like block chain is missing in the implemented measures(Xiong and Xiong 2020; Liang et al. 2020; Kannan et al. 2020), but yet the results reveal the depth and breadth of security strategies employed in different research articles. It provides a comprehensive overview of how various security measures are interlinked and offers insights into potential areas for further investigation.

Table 8 Anomaly detection algorithms

Anomaly category	Ref.	Algorithms	type	Description
Electric power generation and distribution	(Shapsough et al. 2020) (Pei et al. 2022) (Abedi et al. 2023) (Yip et al. 2018b) (Kirbaş and Kerem 2021) (Mondal et al. 2019)	Siamese neural network k-Nearest Neighbors (kNN) Improved Support Vector Machine (SVM) Random Forest Seasonal Autoregressive Moving Average (SARMA) Logistic Regression (LR) Hidden Markov Modeling (HMM) multilayer perceptron	Supervised	Siamese neural networks learn similarity measures to detect unusual patterns. kNN compares new data to known sets, flagging deviations. SVM and Random Forests use advanced modeling and ensemble learning to identify anomalies. SARMA handles time-series data by accounting for seasonal variations.HMM estimate sequence probabilities to detect anomalies, and MLP leverage deep learning to learn complex patterns.
Electric power generation and distribution	(Jiang et al. 2022a) (Leng and Qiu 2023) (al Rashid et al. 2022) (Shapsough et al. 2020)	Isolation Forest algorithm Fuzzy C-means algorithm CNN-LSTM based auto-encoder Latent space NN	Unsupervised	Isolation Forest isolates anomalies by partitioning data, Fuzzy C-means clusters data with varying membership degrees, CNN-LSTM auto-encoders capture spatial-temporal patterns to flag deviations, and Latent space neural networks detect anomalies by learning compact data representations.
Road Lighting and Distribution Systems	(Śmiałkowski and Czyżewski 2022)	LSTM	Unsupervised	Continuously feeding new data and flagging anomalies as they occur.
Battery Faults and Performance	(Jiang et al. 2022a)	Isolated forest algorithm	Unsupervised	After training, establish a threshold for anomaly scores to distinguish between normal and anomalous data points. Points with scores above this threshold are considered anomalies.
Water Quality and Distribution	(Cheng et al. 2023)	Local Outlier Factor (LOF) algorithm	Unsupervised	Calculates the local density of each data point and compares it to the densities of its neighbors. Points that have a significantly lower density than their neighbors are considered outliers.
Resource Extraction and Transportation	(Ramba et al. 2021) (Cheng et al. 2023)	N/A N/A (the results from paper screening are just mentioning ML in general)		

Table 8 (continued)

Anomaly category	Ref.	Algorithms	type	Description
Industrial Production and Maintenance	(Guillen et al. 2020) (Chen et al. 2022b) (Jiang and Zhao 2022)	Long Short-Term Memory (LSTM) and RNN Neural network with adaptive threshold Attention classification-and-segmentation network	Supervised	Capturing temporal dependencies, dynamically adjusting sensitivity, and localizing anomalies within complex data.
Industrial Production and Maintenance	(Chakraborty et al. 2008)	Symbolic Dynamic Filtering (SDF)	Unsupervised	Uncovers hidden patterns and anomalies without relying on predefined labels.
Monitoring and Security	(Azhar et al. 2022)	CNN, LSTM	Supervised	Spatial and temporal patterns in complex data
Monitoring and Security	(Pandey et al. 2020)	DBSCAN, UKF	Unsupervised	Clustering transaction data and identifying unusual patterns or behaviors that deviate from normal customer activity, signaling potential fraudulent activity.
Energy Consumption	(Sawas and Farag 2023)	LSTM	Supervised	Continuously feeding new data and flagging anomalies as they occur.
Energy Consumption	(Yin et al. 2022) (Park et al. 2023) (Alsalemi et al. 2022)	Rain Flow-based Connectivity Outlier Factor algorithm GCNNs M2SP-EdgeIoE	Unsupervised	The Rain Flow algorithm detects irregularities in cyclical patterns. GCNNs effectively identify anomalies in graph-structured data. M2SP-EdgeIoE enables real-time anomaly detection in IoT environments by integrating multi-sensor data and edge computing.
HVAC and Conditioning Systems	(Chou and Telaga 2014b)	Hybrid neural net ARIMA / two-sigma rule	Supervised	Used for time series data, combining statistical modeling with machine learning to identify unusual patterns and deviations.

Table 8 (continued)

Anomaly category	Ref.	Algorithms	type	Description
Renewable Energy	(Mallor et al. 2017) (Shapsough et al. 2020) (Mak-Hau et al. 2022) (Pandit and Infield 2018)	Siamese neural network, NN Isolation Forest algorithm, Latent space, NN, kNN	Supervised Unsupervised	The Isolation Forest algorithm identifies potential outliers, then Siamese neural networks assess the similarity of these outliers to known patterns. Latent space representations are used to reduce the dimensionality of the data, aiding in visualization and subsequent analysis. Neural networks are used to detect complex patterns, while kNN provides density-based estimations to confirm the anomalous nature of the identified data points.

Machine learning: integration and applications

Understanding the AI algorithms behind anomaly detection in research papers allows us to grasp their strengths and weaknesses, which is crucial for choosing the right tool for the job. This knowledge can also help improve existing algorithms by identifying areas for development and customization.

The results of this study provides a comprehensive overview of various algorithms used for anomaly detection, their features, classification (supervised or unsupervised). These information along with a brief description of how these algorithms function and associated references are reported in Table 8. The categories cover a wide range of subsectors in SES, and the algorithms listed, such as Siamese neural networks, k-Nearest Neighbors (kNN), Support Vector Machine (SVM), and Long Short-Term Memory (LSTM), showcase the diversity of techniques employed to detect anomalies.

The results of investigating the implement algorithms provide insights into the mechanisms of each algorithm, highlighting their unique approaches to identifying deviations from normal patterns.

Enhancing anomaly detection performance

Recent studies have demonstrated various effective methods for improving anomaly detection in complex energy systems. This section explores key approaches through empirical evidence and real-world implementations.

Data preprocessing and feature engineering

Hu et al. (2022) conducted a study on smart grid data from 10,000 sensors, showing that proper feature engineering improved anomaly detection accuracy by 18%. They used Principal Component Analysis (PCA) to reduce 500 initial features to 50, resulting in a 30% reduction in computational time without sacrificing accuracy.

In a study of industrial IoT data, Li et al. (2022) applied Independent Component Analysis (ICA) to separate mixed signals from 1,000 sensors. This preprocessing step increased the precision of anomaly detection by 22% compared to using raw sensor data.

Xu et al. (2021) demonstrated the importance of data cleaning in a smart building study. By addressing missing values and outliers in a dataset of 5 million readings from 500 sensors, they reduced false positives in anomaly detection by 35%.

Advanced machine learning and hybrid models

Wang et al. (2023b) employed Bidirectional Long Short-Term Memory (Bi-LSTM) networks in an industrial control system, achieving 98.7% accuracy in anomaly detection on a dataset of 100,000 data points from a simulated water treatment plant.

In cybersecurity, Azhar et al. (2022) compared neural network architectures using the NSL-KDD dataset. Their results showed Gated Recurrent Units (GRUs) outperforming other models with 97.2% accuracy in detecting network intrusions.

Alsalemi et al. (2022) developed a hybrid model integrating statistical methods with neural networks for smart building anomaly detection. Testing on 1 million sensor readings, their approach reduced false positives by 15% compared to standalone neural networks.

Chahla et al. (2020) explored hyperdimensional computing (HDC) for IoT security, demonstrating a 30% reduction in computational resources while maintaining accuracy comparable to deep learning models on a dataset of 50,000 IoT device interactions.

Real-time processing and adaptive methods

Yen et al. (2019) showed a 25% improvement in anomaly detection accuracy by increasing smart meter data collection frequency from hourly to 15-minute intervals in a study of 1000 residential meters over 6 months.

Tehrani et al. (2022) implemented adaptive thresholds in an industrial IoT setting with 500 sensors, reducing false positives by 20% compared to static thresholds.

Jiang et al. (2022a) applied transfer learning to wind turbine anomaly detection, reducing training time for new turbine models by 60% while maintaining detection accuracy.

Signal decomposition and statistical techniques

Jiang et al. (2022a) used Variational Mode Decomposition (VMD) on smart grid power quality data, improving anomaly detection accuracy by 22% compared to methods using raw signals.

Veerakumar et al. (2023) employed Bayesian networks for anomaly detection in a 10,000-node smart grid, achieving a 12% higher F1-score in identifying cyber-attacks compared to traditional threshold-based methods.

Yin et al. (2023) utilized Markov models to detect anomalies in energy consumption patterns of 5,000 households, identifying unusual behaviors with 93% accuracy, a 15% improvement over simple statistical approaches.

Context-specific and domain knowledge integration

Shahid et al. (2023) developed a context-specific anomaly detection system for a solar power plant, integrating expert knowledge about panel degradation and weather effects.

This approach improved detection accuracy by 28% compared to generic machine learning models.

Wong et al. (2021) demonstrated the importance of sensor calibration in a study of 200 urban air quality sensors. Their AI-driven calibration system maintained 95% accuracy over a year, compared to 75% for non-calibrated sensors.

Li et al. (2020) highlighted the significance of expert judgment in anomaly definition for nuclear power plant operations. Involving domain experts in initial data labeling improved their machine learning model's accuracy by 20% compared to purely statistical definitions.

These case studies demonstrate the effectiveness of various techniques in enhancing anomaly detection performance. From advanced data preprocessing and machine learning models to real-time adaptive methods and domain knowledge integration, these approaches significantly improve accuracy, reduce false positives, and adapt to dynamic conditions in complex energy systems. However, they also underscore the need for careful implementation, continuous refinement, and validation by domain experts to achieve optimal results in real-world scenarios.

Standard tools for anomaly detection

Machine learning and statistical methods for anomaly detection

Abnormality detection is widely used by machine learning algorithms because they can handle large and complex datasets. Isolated Forest is a common algorithm for identifying anomalies by isolating observations from a dataset (Hu et al. 2022; Li et al. 2022; Chen et al. 2022b; Ramesh et al. 2022; Shapsough et al. 2020), and Support Vector Machines (SVM), effective for fault classification and anomaly detection in various applications (Tehrani et al. 2022; Pei et al. 2022; Yip et al. 2018b; Chou and Telaga 2014b). K-means and other clustering algorithms are used for grouping similar data points and detecting outliers (Hu et al. 2022; Tehrani et al. 2022; Pandey et al. 2020; Chahla et al. 2020). Long Short-Term Memory (LSTM) networks are applied for time series analysis and anomaly detection (Li et al. 2022; Baker and Shadmand 2023; Sawas and Farag 2023; al Rashid et al. 2022; Azhar et al. 2022; Xu et al. 2023). Convolutional Neural Networks (CNN) are utilized for feature extraction and pattern recognition in images and time-series data (Jiang et al. 2022a; Azhar et al. 2022).

Traditional statistical methods are also prevalent, often used in combination with machine learning techniques. Principal Component Analysis (PCA) is employed for dimensionality reduction and anomaly detection (Bhaskar et al. 2023; Chen et al. 2022a; Wadi and Elmasry 2021). Kalman Filters are applied for state estimation and filtering noise from data (Veerakumar et al. 2023; Xu et al. 2023; Ren et al. 2018). The chi-square test is used for model-based anomaly detection (Abedi et al. 2015).

Hybrid, ensemble models and real-time processing

Combining multiple methods can enhance detection performance. Hybrid models, such as combining CNN and LSTM for stability detection (Azhar et al. 2022), or using ARIMA and ANN for time series data (Chou and Telaga 2014b), are effective. Ensemble methods leverage the strengths of different models to improve robustness and accuracy (Śmiałkowski and Czyżewski 2022; Chen et al. 2022b; Wang et al. 2023b; Chen et al.

2022a; Liu and Aldrich 2023). Despite the fact that hybrid methods have more complexity and require more processing time, they still play a crucial role in real-time anomaly detection.

Advanced tools and platforms facilitate the implementation and scaling of anomaly detection systems. Apache Spark and Flink are used for distributed processing and real-time data analysis (Abedi et al. 2023; Pei et al. 2022). GridLab-D software is utilized for simulating power systems and analyzing anomalies (Sawas and Farag 2023). IoT devices and sensors enable real-time data collection and monitoring, integrated with machine learning algorithms for anomaly detection (Wong et al. 2021; Ramba et al. 2021).

Deep learning techniques provide powerful tools for handling complex data patterns. Autoencoders are used for learning representations and detecting anomalies in various data types (Abedi et al. 2023; Chen et al. 2022b; Yan et al. 2023; Shahid et al. 2023). Deep Neural Networks (DNN), including Transformer models and other architectures, are employed for sophisticated anomaly detection tasks (Chen et al. 2022b; Wang et al. 2023b; Chen et al. 2023; Yan et al. 2023).

Real-time processing capabilities are essential for timely anomaly detection. Real-time monitoring systems are implemented using platforms like Apache Kafka and Spark Streaming for continuous data flow and analysis (Pei et al. 2022). Big data analytics utilizes large datasets for training and improving anomaly detection models, often through distributed systems (Abedi et al. 2023; Li et al. 2020).

Certain methodologies are tailored for specific applications within anomaly detection. Rain Flow-based algorithms are used for point and collective anomaly detection in specific contexts (Yin et al. 2022). Tomographic Inversion Algorithms are applied in detecting anomalies in power systems (Yan et al. 2023). Motif-based association rule mining is used in analyzing smart meter datasets for pattern recognition and anomaly detection (Funde et al. 2019).

These tools and methodologies are used in literature for anomaly detection in complex energy systems. By integrating machine learning algorithms, statistical methods, hybrid models, advanced platforms, and real-time processing capabilities, robust and reliable anomaly detection systems can be developed. These methods ensure the systems are adaptable, accurate, and efficient in dynamic and complex environments.

Despite significant advancements in anomaly detection methodologies and tools, several gaps need to be addressed to improve the reliability and practical implementation of these systems in complex energy environments. One of the primary gaps is the lack of evaluation and deployment of advanced anomaly detection techniques in real energy system factories. Many studies focus on theoretical developments and simulations without validating their findings in practical, real-world settings. This gap highlights the need for extensive field trials and industrial collaborations to ensure that these advanced methodologies can effectively handle the complexities and dynamic nature of actual energy systems.

Furthermore, there is a noticeable dependency on cutting-edge technologies which, although promising, require time to mature and become reliable through extensive testing and iterative refinement. This maturity process is often overlooked,

leading to premature implementation and potential inefficiencies. It is crucial to adopt a cautious approach, starting with offline deployments to validate new algorithms before integrating them fully into live systems.

Addressing these gaps through comprehensive evaluations, iterative development, and close collaboration between academia and industry will be crucial for advancing the field of anomaly detection in complex energy systems.

Enabling technologies for real-time anomaly detection

Recent advancements have significantly enhanced real-time anomaly detection across various applications. This section explores key enabling technologies through empirical studies and real-world implementations.

Sensor technologies and data acquisition

Advanced sensor technologies and IoT devices form the foundation of real-time data collection and monitoring. Veerakumar et al. (2023) demonstrated the efficacy of high-speed Phasor Measurement Units (PMUs) in a smart grid pilot project. Their study, involving 50 PMUs across a regional power network, achieved a 99.9% accuracy in detecting voltage anomalies within 100 milliseconds, a 40% improvement over traditional methods.

In the realm of smart metering, Ramba et al. (2021) conducted a large-scale deployment of Advanced Metering Infrastructure (AMI) in a city of 500,000 residents. The AMI system, collecting data at 15-minute intervals, enabled the detection of water leaks within 2 hours of occurrence, reducing water loss by 30% over a one-year period compared to previous systems.

Data processing and machine learning algorithms

Machine learning algorithms have revolutionized anomaly detection capabilities. Sawas and Farag (2023) applied Long Short-Term Memory (LSTM) networks to network traffic data from a major telecommunications provider. Their model, processing 1 million packets per second, achieved a 96% accuracy in identifying cyber attacks, a 15% improvement over traditional signature-based methods.

Chen et al. (2022b) utilized Convolutional Neural Networks (CNNs) for anomaly detection in industrial IoT settings. In a factory with 1000 sensors, their CNN model reduced false positives by 40% compared to threshold-based approaches, while maintaining a 98% true positive rate.

Edge and cloud computing

The integration of edge and cloud computing has significantly enhanced real-time processing capabilities. Liu and Aldrich (2023) implemented an edge computing solution in a smart manufacturing plant. By processing data from 500 sensors locally, they reduced response time to anomalies from 5 seconds to 200 milliseconds, crucial for preventing equipment failures.

Hu et al. (2022) leveraged cloud computing for a city-wide traffic anomaly detection system. Their cloud-based platform, processing data from 10,000 traffic sensors, scaled

to handle a 500% increase in data volume during peak hours while maintaining a consistent anomaly detection latency of under 1 s.

Advanced machine learning frameworks

Chen et al. (2023) demonstrated the potential of TinyML in resource-constrained environments. They deployed TinyML models on 1000 low-cost sensors in a smart agriculture setting. Despite having 100 times fewer parameters than cloud-based models, these edge devices achieved 92% accuracy in detecting crop diseases, with a 60% reduction in power consumption.

Jiang et al. (2022a) applied transfer learning techniques to anomaly detection in wind turbines. By transferring knowledge from models trained on 50 older turbines to 10 new turbine models, they reduced the required training data by 70% while maintaining a 95% detection accuracy for mechanical faults.

Communication networks and data transmission

Advancements in communication technologies have been crucial for real-time data transmission. Shapsough et al. (2020) implemented an MQTT-based communication system in a smart building with 5000 IoT devices. Their approach reduced data transmission latency by 65% compared to HTTP, enabling real-time anomaly detection in energy consumption patterns.

Hong et al. (2011) developed a Distributed Intrusion Detection System (DIDS) for a network of 1000 IoT devices. By enabling devices to share threat information in real-time, they improved overall anomaly detection accuracy by 25% compared to isolated device detection.

Visualization and user interface

Effective visualization tools have proven essential for interpreting anomaly detection results. Yin et al. (2023) developed a real-time visualization platform for a power grid control center, monitoring 100,000 nodes. Their interface reduced operator response time to critical anomalies by 40%, from an average of 5 to 3 min, significantly improving grid stability during fault conditions.

These case studies illustrate the transformative impact of enabling technologies on real-time anomaly detection. From advanced sensors and sophisticated algorithms to edge-cloud integration and improved communication networks, these technologies have dramatically enhanced the speed, accuracy, and scalability of anomaly detection systems across diverse applications. However, the studies also highlight ongoing challenges, such as managing increasing data volumes and ensuring system reliability in dynamic environments, pointing to areas for future research and development.

Advancements and future research lines in real-time anomaly detection

Advancements and future studies in real-time anomaly detection are essential to enhance the accuracy, efficiency, and reliability of these systems in complex environments. This section synthesizes the key areas of improvement based on the literature.

Improving data quality and sensor accuracy

Improving data quality and sensor accuracy is fundamental for enhancing anomaly detection. Strategies include using accurate sensors, mitigating the impact of measurement errors, and exploring statistical testing with multiple cell anomalies Hu et al. (2022); Shoman and Burr (2023); Bhaskar et al. (2023). Enhanced data cleaning and dimensionality reduction techniques are also crucial for better detection performance Chen et al. (2022b).

Optimization and integration of advanced technologies

Optimizing fault diagnosis methods for real-time efficiency is necessary Jiang et al. (2022a). Integrating edge computing, enhancing data fusion, and developing adaptive algorithms will improve system maintenance and anomaly detection Śmiałkowski and Czyżewski (2022). Future studies should explore robust Hyperdimensional Computing (HDC) methods Wang et al. (2022) and advanced machine learning algorithms such as deep learning, LSTM networks, and hybrid models combining CNN and LSTM Tehrani et al. (2022); Sawas and Farag (2023); al Rashid et al. (2022); Azhar et al. (2022); Chen et al. (2022a). Additionally, improving anomaly detection capabilities against challenging scenarios like malicious data attacks and topology changes is crucial Veerakumar et al. (2023); al Rashid et al. (2022).

Developing robust and adaptive algorithms

Future research should focus on developing robust and adaptive algorithms that can handle dynamic environments and diverse data sources. This includes implementing multi-fault diagnosis, considering environmental factors, enhancing model-free/data-driven methods, and further developing fault prediction methods using deep learning techniques Sun et al. (2022). The creation of a unified framework with data-efficient training, automatic adjustment mechanisms, and reduced computational costs will also be beneficial Wang et al. (2023b).

Enhancing real-time processing and scalability

Advancements in real-time anomaly detection should aim to reduce the Mean Time to Detection (MTTD), improve model enhancements, and estimate onset time Sawas and Farag (2023). Optimizing real-time model training, integrating anomaly detection models with manufacturing processes, and using advanced clustering algorithms like K-means are necessary for better performance Yan et al. (2023); Chahla et al. (2020). Implementing big data infrastructure for accommodating and analyzing unstructured data from smart grids, and incorporating optimization algorithms and explanatory variables into building management systems will enhance anomaly detection Chou and Telaga (2014b).

Ensuring security and resilience

Improving the resilience of anomaly detection systems to false alarms and cyber threats is crucial Hong et al. (2011). Developing more sophisticated algorithms, enhancing coordination between distributed agents in a network, and incorporating security measures to protect against cyber threats are essential Abedi et al. (2023);

Hong et al. (2011). Additionally, future studies should focus on improving the accuracy and reliability of intrusion detection methods and exploring advanced ML models for anomaly detection Abedi et al. (2015).

Integrating domain knowledge and adaptive control strategies

Integrating domain knowledge into anomaly detection algorithms, exploring novel data sources and feature extraction techniques, and researching automated model retraining and adaptation are critical areas for future research Guillen et al. (2020); Pandey et al. (2020). Improving the accuracy and reliability of sensor data, enhancing the robustness of anomaly detection algorithms against false positives, and developing adaptive control strategies that can dynamically adjust system parameters based on real-time conditions will improve the accuracy and reliability of anomaly detection systems Li et al. (2020); Pandey et al. (2020).

This synthesis highlights the diverse array of advancements and future studies needed in real-time anomaly detection to ensure accurate functioning. By focusing on improving data quality, optimizing advanced technologies, developing robust algorithms, enhancing real-time processing, ensuring security, and integrating domain knowledge, anomaly detection systems can become more effective and reliable in complex environments.

Conclusion

This systematic literature review examined real-time anomaly detection within complex energy systems, revealing critical trends, methodologies, and essential security measures. The review methodology encompassed a mixed-methods approach, combining quantitative and qualitative analyses, and a longitudinal design to track developments over time. By rigorously evaluating 79 articles through descriptive, exploratory, and comparative studies, we aimed to provide a comprehensive understanding of the state-of-the-art, challenges, and opportunities in this field.

Our findings highlight the growing reliance on machine learning and deep learning techniques to identify novel and sophisticated cyber threats for anomaly detection in Energy Systems (Aghazadeh Ardebili et al. 2024). However, the successful deployment of these systems necessitates robust security measures to protect data integrity and confidentiality. Key security measures identified include data encryption, access control, intrusion detection, and deployment in trusted environments. The network graph of the articles in the Sect. "Security of anomaly detection ecosystem" illustrates the relationships between the methods that are employed. The results show that data encryption and access control are widely implemented security measures. However, significant research gaps exist in areas such as resource assessment and compliance with standards. Addressing these gaps is crucial for developing more comprehensive and resilient anomaly detection systems.

In conclusion, real-time anomaly detection is integral to the security and efficiency of complex energy systems. Implementing diverse and layered security strategies is essential to safeguard these systems from evolving cyber threats. Future research should focus on underexplored areas, such as resource assessment and blockchain integration, to further enhance system robustness and reliability.

Nevertheless, it is important at this point to respect the intrinsic limitations incorporated in this study: The selection criteria, as well as subjective interpretation of findings, one should consider, so some bias might be introduced in this work. Another point to consider is that this area is rapidly developing; such recent achievements in technology for anomaly detection might not be fully described in the literature. Non-uniformity of studies conducted based on the field of anomaly detection through AI, the combination with inconsistent KPIs, and differences in the design studies themselves might altogether be contributing to bias in the quality and heterogeneity of the included studies.

The review, therefore, focuses mainly on peculiar features of resilience and anomaly detection properties of the energy systems as configured in the selection phase presented in Sect. "Conducting phase". This does, presumably, have the result of excluding wider or cross-disciplinary approaches which might provide valuable insights and so may limit the generalizability of findings to other contexts. Similarly, for technical disciplines, there is potential for publication bias because grey literature, including important material such as technical reports coming from industry, is not included; these reports are not usually peer-reviewed or formally published.

Acknowledgements
The research was partially supported by HSPI SpA and the Italian Research Center on High Performance Computing, Big Data and Quantum Computing (ICSC) grant, funded by EU - NextGenerationEU (PNRR-HPC, CUP:C83C22000560007), and the RIPARTI regional project - dataEnrichment for Resilient UAS (assegni di RIcerca per riPARTire con le Imprese)-POC PUGLIA FESRTFSE 2014/2020,CUP F87G22000270002.

Author contributions
AAA, OH, and AB contributed in providing substantial input in the conceptualization and methodology. Authors actively contributed throughout the research analysis, validation and writing.

Funding
Not applicable.

Availability of data and materials
Not applicable

Code availability
Not applicable

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors have no relevant financial or non-financial interests to disclose.

Received: 15 June 2024 Accepted: 15 September 2024
Published online: 04 October 2024

References

Abdelmoula IA, Elhamaoui S, Mehdary A, et al (2023) Application of a data-driven anomaly detection approach for a solar photovoltaic plant using real-time scada data. In: 2023 International Conference on Digital Age & Technological Advances for Sustainable Development (ICDATA), IEEE, pp 106–112

Abedi S, Arvani A, Jamalzadeh R (2015) Cyber Security of Plug-in Electric Vehicles in Smart Grids: Application of Intrusion Detection Methods, pp 129–147. https://doi.org/10.1007/978-981-287-299-9_5

Abedi A, Rajkumar VS, Ştefanov A, et al (2023) Towards real-time distinction of power system faults and cyber attacks. IEEE, pp 1–5, https://doi.org/10.1109/PESGM52003.2023.10253241

Aghazadeh Ardebili A, Martella C, Martella A, et al (2024) Smart critical infrastructures security management and governance: Implementation of cyber resilience kpis for decentralized energy asset. In: CEUR Workshop Proceedings. https://ceur-ws.org/Vol-3731/paper24.pdf

Albogamy FR, Paracha MYI, Hafeez G et al (2022) Real-time scheduling for optimal energy optimization in smart grid integrated with renewable energy sources. IEEE Access 10:35498–35520. https://doi.org/10.1109/ACCESS.2022.3161845
Alli-Balogun M (2024) Application of text summarization on text-based generative adversarial networks. Int J Comput (IJC) 50(1):8–31
al Rashid AM, Hossain F, Anwar A, et al (2022) False data injection attack detection in smart grid using energy consumption forecasting. Energies 15:4877. https://doi.org/10.3390/en15134877
Alsalemi A, Himeur Y, Bensaali F et al (2022) An innovative edge-based internet of energy solution for promoting energy saving in buildings. Sustain Cities Soc 78:103571. https://doi.org/10.1016/j.scs.2021.103571
Anagnostou G, Boem F, Kuenzel S et al (2018) Observer-based anomaly detection of synchronous generators for power systems monitoring. IEEE Trans Power Syst 33:4228–4237. https://doi.org/10.1109/TPWRS.2017.2771278
Ayaz E, Şeker S, Barutçu B, et al (2003) Comparisons between the various types of neural networks with the data of wide range operational conditions of the borssele npp. Prog Nucl Energy 43:381–387. https://doi.org/10.1016/S0149-1970(03)00047-7
Azhar IF, Putranto LM, Irnawan R (2022) Development of pmu-based transient stability detection methods using CNN-LSTM considering time series data measurement. Energies 15:8241. https://doi.org/10.3390/en15218241
Baker M, Shadmand MB (2023) An lstm-based anomaly classification framework for power electronics dominated grids. IEEE, pp 1–7. https://doi.org/10.1109/PECI57361.2023.10197777
Bhaskar K, Kumar A, Bunce J et al (2023) Data-driven thermal anomaly detection in large battery packs. Batteries 9:70. https://doi.org/10.3390/batteries9020070
Brahma S, Kavasseri R, Cao H et al (2016) Real-time identification of dynamic events in power systems using pmu data, and potential applications-models, promises, and challenges. IEEE Trans Power Deliv 32(1):294–301
Bushehri AS, Keivanpour S, Azam M, et al (2021) Anomalous energy detection for resource-constrained embedded systems using tracing data analysis. IEEE, pp 1–8, https://doi.org/10.1109/ICECET52533.2021.9698794
Cadini F, Sbarufatti C, Cancelliere F et al (2019) State-of-life prognosis and diagnosis of lithium-ion batteries by data-driven particle filters. Appl Energy 235:661–672. https://doi.org/10.1016/j.apenergy.2018.10.095
Capozzoli A, Piscitelli MS, Brandi S et al (2018) Automated load pattern learning and anomaly detection for enhancing energy management in smart buildings. Energy 157:336–352
Chahla C, Snoussi H, Merghem L et al (2020) A deep learning approach for anomaly detection and prediction in power consumption data. Energy Effic13:1633–1651. https://doi.org/10.1007/s12053-020-09884-2
Chakraborty S, Sarkar S, Gupta S et al (2008) Damage monitoring of refractory wall in a generic entrained-bed slagging gasification system. Proc Inst Mech Eng Part A J Power Energy 222:791–807. https://doi.org/10.1243/09576509JPE638
Chen J, Liu ZS, Jiang H et al (2022) Anomaly detection of control rod drive mechanism using long short-term memory-based autoencoder and extreme gradient boosting. Nucl Sci Tech 33:127. https://doi.org/10.1007/s41365-022-01111-0
Chen Y, Huang Y, Miao B et al (2022) Adaptive anomaly detection-based liquid loading prediction in shale gas wells. J Petrol Sci Eng 214:110522. https://doi.org/10.1016/j.petrol.2022.110522
Chen Z, Gao Y, Liang J (2023) A self-powered sensing system with embedded tinyml for anomaly detection. IEEE, pp 1–6, https://doi.org/10.1109/IESES53571.2023.10253705
Cheng M, Zhang D, Yan W et al (2023) Power system abnormal pattern detection for new energy big data. Int J Emerg Electr Power Syst 24(1):91–102
Chou JS, Telaga AS (2014) Real-time detection of anomalous power consumption. Renew Sustain Energy Revs 33:400–411
Chou JS, Telaga AS (2014) Real-time detection of anomalous power consumption. Renew Sustain Energy Rev 33:400–411. https://doi.org/10.1016/j.rser.2014.01.088
Colak I, Sagiroglu S, Fulli G et al (2016) A survey on the critical issues in smart grid technologies. Renew Sustain Energy Rev 54:396–405. https://doi.org/10.1016/j.rser.2015.10.036
Davarifar M, Rabhi A, Hajjaji A, et al (2014) Real-time diagnosis of pv system by using the sequential probability ratio test (sprt). IEEE, pp 508–513, https://doi.org/10.1109/EPEPEMC.2014.6980544
Eichler C, Röckl J, Jung B, et al (2024) Profiling with trust: system monitoring from trusted execution environments. Design Automation for Embedded Systems pp 1–22
Ekti AR, Wilson A, Olatt J et al (2022) A simple and accurate energy-detector-based transient waveform detection for smart grids: Real-world field data performance. Energies 15:8367. https://doi.org/10.3390/en15228367
Funde NA, Dhabu MM, Paramasivam A et al (2019) Motif-based association rule mining and clustering technique for determining energy usage patterns for smart meter data. Sustain Cities Soc 46:101415. https://doi.org/10.1016/j.scs.2018.12.043
Giunta G, Nielsen KL, Bernasconi G et al (2019) Data driven smart monitoring for pipeline integrity assessment. SPE. https://doi.org/10.2118/197327-MS
Goodhue DL, Quillard JA, Rockart JF (1988) Managing the data resource: a contingency perspective. MIS Q 373–392
Graber R, Kleinhammer R (2016) Constructing the "best" reliability data for the job. In: 2016 Annual Reliability and Maintainability Symposium (RAMS), IEEE, pp 1–6
Guillen D, Anderson N, Krome C et al (2020) A relap5-3d/lstm model for the analysis of drywell cooling fan failure. Prog Nucl Energy 130:103540. https://doi.org/10.1016/j.pnucene.2020.103540
Haq EU, Pei C, Zhang R et al (2023) Electricity-theft detection for smart grid security using smart meter data: a deep-cnn based approach. Energy Rep 9:634–643. https://doi.org/10.1016/j.egyr.2022.11.072
Hong J, Wu SS, Stefanov A, et al (2011) An intrusion and defense testbed in a cyber-power system environment. IEEE, pp 1–5, https://doi.org/10.1109/PES.2011.6039375
Hu Z, Chen W, Wang H et al (2022) Integrated data-driven framework for anomaly detection and early warning in water distribution system. J Clean Prod 373:133977. https://doi.org/10.1016/j.jclepro.2022.133977

Ji X, Yin Z, Zhang Y et al (2021) Real-time robust forecasting-aided state estimation of power system based on data-driven models. Int J Electr Power Energy Syst 125:106412
Jiang Y, Zhao C (2022) Attention classification-and-segmentation network for micro-crack anomaly detection of photovoltaic module cells. Solar Energy 238:291–304. https://doi.org/10.1016/j.solener.2022.04.012
Jiang J, Li T, Chang C et al (2022) Fault diagnosis method for lithium-ion batteries in electric vehicles based on isolated forest algorithm. J Energy Storage 50:104177. https://doi.org/10.1016/j.est.2022.104177
Jiang Y, Dong X, Xu A, et al (2022b) A multi-agent based wide-area protection scheme for distributed cyber energy system. IEEE, pp 1531–1537, https://doi.org/10.1109/EI256261.2022.10116819
Jin X, Guo Y, Sarkar S et al (2011) Anomaly detection in nuclear power plants via symbolic dynamic filtering. IEEE Trans Nucl Sci 58:277–288. https://doi.org/10.1109/TNS.2010.2088138
Jin X, Sun Y, Que Z et al (2016) Anomaly detection and fault prognosis for bearings. IEEE Trans Instrum Meas 65(9):2046–2054
Kamat P, Sugandhi R (2020) Anomaly detection for predictive maintenance in industry 4.0-a survey. In: E3S web of conferences, EDP Sciences, p 02007
Kannan K, Singh A, Verma M, et al (2020) Blockchain-based platform for trusted collaborations on data and ai models. In: 2020 IEEE International Conference on Blockchain (Blockchain), IEEE, pp 82–89
Khediri A, Laouar MR (2018) Deep-belief network based prediction model for power outage in smart grid. ACM, pp 1–6, https://doi.org/10.1145/3213187.3287611
Kirbaş I, Kerem A (2021) A new vibration-based hybrid anomaly detection model for preventing high-power generator failures in power plants. Energy Sources Part A Recov Util Environ Effects 43:3184–3202. https://doi.org/10.1080/15567036.2021.1960654
Klaes M, Narayan A, Patil AD et al (2020) State description of cyber-physical energy systems. Energy Inform 3:1–19
Krishna VB, Jung D, Khiem NQM, et al (2013) Energytrack. ACM, pp 1–2, https://doi.org/10.1145/2528282.2534158
Leal-Arcas R, Boskovic S, Karimabadi MSA (2020) The transition to decentralized energy: challenges, opportunities and progress. UPR Bus LJ 11:1
Leng D, Qiu Z (2023) Identification of anomaly detection in power system state estimation based on fuzzy c-means algorithm. Int Trans Electric Energy Syst 2023:1–12. https://doi.org/10.1155/2023/7553080
Li W, Koo C, Hong T et al (2020) A novel operation approach for the energy efficiency improvement of the hvac system in office spaces through real-time big data analytics. Renew Sustain Energy Rev 127:109885. https://doi.org/10.1016/j.rser.2020.109885
Li S, Pandey A, Hooi B et al (2022) Dynamic graph-based anomaly detection in the electrical grid. IEEE Trans Power Syst 37:3408–3422. https://doi.org/10.1109/TPWRS.2021.3132852
Liang S, Baozhong H, Yang L, et al (2020) Blockchain-based power grid data asset management architecture. In: 2020 International Conference on Computer Science and Management Technology (ICCSMT), IEEE, pp 207–211
Liu X, Aldrich C (2023) Explaining anomalies in coal proximity and coal processing data with shapley and tree-based models. Fuel 335:126891. https://doi.org/10.1016/j.fuel.2022.126891
Mak-Hau V, Henkel A, Abdelrazek M, et al (2022) Ddt: the deakin university microgrid digital twin. IEEE, pp 1–6, https://doi.org/10.1109/APPEEC53445.2022.10072197
Mallor F, León T, Boeck LD, et al (2017) A method for detecting malfunctions in pv solar panels based on electricity production monitoring. Solar Energy 153:51–63. https://doi.org/10.1016/j.solener.2017.05.014
Mazumder SK, Shadmand M, Mantooth HA, et al (2024) Power grid resilience. In: Power electronics handbook. Elsevier, p 1015–1033
Mead DA (1982) Evaluating the quality of data used for resource planning. Proc of Working Party S 3
Mihalcea R, Radev D (2011) Graph-based natural language processing and information retrieval. Cambridge: Cambridge University Press
Mohammadpourfard M, Genc I, Lakshminarayana S, et al (2021) Attack detection and localization in smart grid with image-based deep learning. IEEE, pp 121–126, https://doi.org/10.1109/SmartGridComm51999.2021.9631994
Mondal S, Ghalyan NF, Ray A et al (2019) Early detection of thermoacoustic instabilities using hidden markov models. Combust Sci Technol 191:1309–1336. https://doi.org/10.1080/00102202.2018.1523900
Narayan A, Brand M, Lehnhoff S (2023) Quantifying the resilience of ict-enabled grid services in cyber-physical energy system. Energy Inform 6(Suppl 1):23
Nguyen AT, Nguyen TN (2015) Graph-based statistical language model for code. In: 2015 IEEE/ACM 37th IEEE International Conference on Software Engineering, IEEE, pp 858–868
Noureen SS, Bayne SB, Shaffer E, et al (2019) Anomaly detection in cyber-physical system using logistic regression analysis. IEEE, pp 1–6, https://doi.org/10.1109/TPEC.2019.8662186
Nur N, Sridhar S, Pal S, et al (2019) A clustering approach for consumer baselining and anomaly detection in transactive control. ACM, pp 516–521, https://doi.org/10.1145/3307772.3331028
Olatunde TM, Okwandu AC, Akande DO et al (2024) The impact of smart grids on energy efficiency: a comprehensive review. Eng Sci Technol J 5:1257–1269. https://doi.org/10.51594/estj.v5i4.1016
Pan H, Yin Z, Jiang X (2022) High-dimensional energy consumption anomaly detection: A deep learning-based method for detecting anomalies. Energies 15:6139. https://doi.org/10.3390/en15176139
Pandey S, Srivastava AK, Amidan BG (2020) A real time event detection, classification and localization using synchrophasor data. IEEE Trans Power Syst 35:4421–4431. https://doi.org/10.1109/TPWRS.2020.2986019
Pandit RK, Infield D (2018) Scada-based wind turbine anomaly detection using gaussian process models for wind turbine condition monitoring purposes. IET Renew Power Gen 12:1249–1255. https://doi.org/10.1049/iet-rpg.2018.0156
Park S, Gama F, Lavaei J, et al (2023) Distributed power system state estimation using graph convolutional neural networks. In: Hawaii International Conference on System Sciences, https://api.semanticscholar.org/CorpusID:243831659
Pei C, Zhang S, Zeng X (2022) Research on anomaly detection of wireless data acquisition in power system based on spark. Energy Rep 8:1392–1404. https://doi.org/10.1016/j.egyr.2022.01.224

Pileggi P, Verriet J, Broekhuijsen J, et al (2019) A digital twin for cyber-physical energy systems. IEEE, pp 1–6, https://doi.org/10.1109/MSCPES.2019.8738792
Pourhabibi T, Ong KL, Kam BH et al (2020) Fraud detection: a systematic literature review of graph-based anomaly detection approaches. Decis Supp Syst 133:113303
Ramba V, Selvaraju S, Muppidathi BV et al (2021) Evaluation of structural integrity of tubulars in directional wellbores: a case study in north-eastern parts of india. J Petrol Sci Eng 207:109067. https://doi.org/10.1016/j.petrol.2021.109067
Ramesh J, Shahriar S, Al-Ali AR et al (2022) Machine learning approach for smart distribution transformers load monitoring and management system. Energies 15:7981. https://doi.org/10.3390/en15217981
Ren H, Hou Z, Etingov P (2018) Online anomaly detection using machine learning and hpc for power system synchrophasor measurements. IEEE, pp 1–5, https://doi.org/10.1109/PMAPS.2018.8440495
Rosch D, Ruhe S, Schafer K, et al (2019) Local anomaly detection analysis in distribution grid based on iec 61850-9-2 le sv voltage signals. IEEE, pp 1–6, https://doi.org/10.1109/SEST.2019.8849139
Sankey ML, Jeter SM, Wolf TD, et al (2014) Continuous monitoring, modeling, and evaluation of actual building energy systems. In: Energy Sustainability, American Society of Mechanical Engineers, p V002T06A006
Sawas A, Farag HE (2023) Real-time detection of stealthy iot-based cyber-attacks on power distribution systems: a novel anomaly prediction approach. Electr Power Syst Res 223:109496. https://doi.org/10.1016/j.epsr.2023.109496
Schäfer M, Kebir N, Neumann K (2011) Research needs for meeting the challenge of decentralized energy supply in developing countries. Energy Sustain Dev 15(3):324–329
Shahid ZK, Saguna S, Ã…hlund C (2023) Autoencoders for anomaly detection in electricity and district heating consumption: a case study in school buildings in sweden. IEEE, pp 1–8, https://doi.org/10.1109/EEEIC/ICPSEurope57605.2023.10194605
Shapsough S, Zualkernan I, Dhaouadi R, et al (2020) Using siamese networks to detect shading on the edge of solar farms. IEEE, pp 1–8, https://doi.org/10.1109/IOTSMS52051.2020.9340189
Shibu NS, Devidas AR, Balamurugan S, et al (2024) Optimising microgrid resilience: Integrating iot, blockchain, and smart contracts for power outage management. IEEE Access
Shoman N, Burr T (2023) Impact of measurement error on deep neural networks for nuclear material accountancy. Nucl Eng Des 402:112113. https://doi.org/10.1016/j.nucengdes.2022.112113
Singh VK, Hossain R, Tucker E (2024) Anomaly detection and mitigation for dynamic frequency regulation in hydropower-battery systems. Tech. rep, National Renewable Energy Laboratory (NREL), Golden, CO (United States)
Sleiti AK, Kapat JS, Vesely L (2022) Digital twin in energy industry: proposed robust digital twin for power plant and other complex capital-intensive large engineering systems. Energy Rep 8:3704–3726
Śmiałkowski T, Czyżewski A (2022) Detection of anomalies in the operation of a road lighting system based on data from smart electricity meters. Energies 15:9438. https://doi.org/10.3390/en15249438
Stewart MK, Stewart C (2024) Massive power system failures. In: Ciottone's Disaster Medicine. Elsevier, p 978–983
Sun J, Qiu Y, Shang Y et al (2022) A multi-fault advanced diagnosis method based on sparse data observers for lithium-ion batteries. J Energy Storage 50:104694. https://doi.org/10.1016/j.est.2022.104694
Tang D, Fang YP, Zio E (2023) Vulnerability analysis of demand-response with renewable energy integration in smart grids to cyber attacks and online detection methods. Reliabil Eng Syst Saf 235:109212. https://doi.org/10.1016/j.ress.2023.109212
Tehrani SO, Shahrestani A, Yaghmaee MH (2022) Online electricity theft detection framework for large-scale smart grid data. Electr Power Syst Res 208:107895. https://doi.org/10.1016/j.epsr.2022.107895
Toshev R (2016) Risks and prospects of smart electric grids systems measured with real options
Urishev B (2019) Decentralized energy systems, based on renewable energy sources. Appl Solar Energy 55:207–212
Van Aken B, Winter B, Löser A, et al (2019) How does bert answer questions? a layer-wise analysis of transformer representations. In: Proceedings of the 28th ACM international conference on information and knowledge management, pp 1823–1832
Veerakumar N, Ćetenović D, Kongurai K, et al (2023) Pmu-based real-time distribution system state estimation considering anomaly detection, discrimination and identification. Int J Electric Power Energy Syst 148:108916. https://doi.org/10.1016/j.ijepes.2022.108916
Vegesna VV (2024) Machine learning approaches for anomaly detection in cyber-physical systems: a case study in critical infrastructure protection. Int J Mach Learn Artif Intell 5(5):1–13
Vikram A, et al (2020) Anomaly detection in network traffic using unsupervised machine learning approach. In: 2020 5th International conference on communication and electronics systems (ICCES), IEEE, pp 476–479
Wadi M, Elmasry W (2021) An anomaly-based technique for fault detection in power system networks. IEEE, pp 1–6, https://doi.org/10.1109/ICEPE-P51568.2021.9423479
Wang Q, Bu S (2020) Deep learning enhanced situation awareness for high renewable-penetrated power systems with multiple data corruptions. IET Renew Power Gener 14:1134–1142. https://doi.org/10.1049/iet-rpg.2019.1015
Wang H, Meng A, Liu Y et al (2019) Unscented kalman filter based interval state estimation of cyber physical energy system for detection of dynamic attack. Energy 188:116036. https://doi.org/10.1016/j.energy.2019.116036
Wang X, Flores R, Brouwer J et al (2022) Real-time detection of electrical load anomalies through hyperdimensional computing. Energy 261:125042. https://doi.org/10.1016/j.energy.2022.125042
Wang W, An A, Zhang Z et al (2023a) Early-warning of generator collusion in chinese electricity market based on information deep autoencoding gaussian mixture model. Electr Power Syst Res 221:109425
Wang X, Yao Z, Papaefthymiou M (2023b) A real-time electrical load forecasting and unsupervised anomaly detection framework. Appl Energy 330:120279. https://doi.org/10.1016/j.apenergy.2022.120279
Wang X, Wang H, Bhandari B et al (2024) Ai-empowered methods for smart energy consumption: a review of load forecasting, anomaly detection and demand response. Int J Precis Eng Manuf-Green Technol 11(3):963–993
Weinand JM, Scheller F, McKenna R (2020) Reviewing energy system modelling of decentralized energy autonomy. Energy 203:117817

Wen W, Liu Y, Sun R et al (2022) Research on anomaly detection of wind farm scada wind speed data. Energies 15:5869. https://doi.org/10.3390/en15165869
Wong YJ, Nakayama R, Shimizu Y, et al (2021) Toward industrial revolution 4.0: Development, validation, and application of 3d-printed iot-based water quality monitoring system. J Clean Prod 324:129230. https://doi.org/10.1016/j.jclepro.2021.129230
Wyss I, Murari A, Spolladore L et al (2023) Comparison of a fast low spatial resolution inversion method and peaking factors for the detection of anomalous radiation patterns and disruption prediction. Fusion Eng Des 193:113625. https://doi.org/10.1016/j.fusengdes.2023.113625
Xiong W, Xiong L (2020) Data resource protection based on smart contract. Comput Secur 98:102004
Xu A, Jiang Y, Cao Y, et al (2019) Addp: Anomaly detection for dtu based on power consumption side-channel. In: 2019 IEEE 3rd Conference on Energy Internet and Energy System Integration (EI2), pp 2659–2663, https://doi.org/10.1109/EI247390.2019.9062014
Xu Y, Yan C, Shi J et al (2021) An anomaly detection and dynamic energy performance evaluation method for hvac systems based on data mining. Sustain Energy Technol Assess 44:101092. https://doi.org/10.1016/j.seta.2021.101092
Xu M, Lu S, Li Z, et al (2023) Research on data anomaly discrimination method for multi-station fusion measurement system. IEEE, pp 1111–1116, https://doi.org/10.1109/AEEES56888.2023.10114246
Yan X, Gao Y, Xu H (2022) Research on power grid anomaly detection based on high-dimensional random matrix theory. IEEE, pp 427–431, https://doi.org/10.1109/IC2ECS57645.2022.10088088
Yan A, Rupnowski P, Guba N et al (2023) Towards deep computer vision for in-line defect detection in polymer electrolyte membrane fuel cell materials. Int J Hydrog Energy 48:18978–18995. https://doi.org/10.1016/j.ijhydene.2023.01.257
Yao Y, Han T, Yu J et al (2024) Uncertainty-aware deep learning for reliable health monitoring in safety-critical energy systems. Energy 291:130419
Yen SW, Morris S, Ezra MA et al (2019) Effect of smart meter data collection frequency in an early detection of shorter-duration voltage anomalies in smart grids. Int J Electric Power Energy Syst 109:1–8. https://doi.org/10.1016/j.ijepes.2019.01.039
Yin S, Yang H, Xu K et al (2022) Dynamic real-time abnormal energy consumption detection and energy efficiency optimization analysis considering uncertainty. Appl Energy 307:118314. https://doi.org/10.1016/j.apenergy.2021.118314
Yin H, Sun K, Chen S, et al (2023) Anomaly detection method for connecting bolts of generator rotor coils based on local enhancement and regional characteristics. IEEE, pp 374–377, https://doi.org/10.1109/ICPECA56706.2023.10075934
Yip SC, Wong K, Hew WP et al (2017) Detection of energy theft and defective smart meters in smart grids using linear regression. Int J Electr Power Energy Syst 91:230–240. https://doi.org/10.1016/j.ijepes.2017.04.005
Yip SC, Tan C, Tan WN, et al (2018a) Detection of energy theft and metering defects in advanced metering infrastructure using analytics. In: 2018 International conference on smart grid and clean energy technologies (ICSGCE), IEEE, pp 15–22
Yip SC, Tan C, Tan WN, et al (2018b) Detection of energy theft and metering defects in advanced metering infrastructure using analytics. IEEE, pp 15–22, https://doi.org/10.1109/ICSGCE.2018.8556710
Zhu J, Xia Y, Wu L, et al (2020) Incorporating bert into neural machine translation. arXiv preprint arXiv:2002.06823

Publisher's Note