
Citation:

Bhukya, R and Moeed, SA and Medavaka, A and Khadidos, AO and Khadidos, AO and Selvarajan, S (2025) SPARK and SAD: Leading-edge deep learning frameworks for robust and effective intrusion detection in SCADA systems. *International Journal of Critical Infrastructure Protection*, 49. pp. 1-24. ISSN 1874-5482 DOI: <https://doi.org/10.1016/j.ijcip.2025.100759>

Link to Leeds Beckett Repository record:

<https://eprints.leedsbeckett.ac.uk/id/eprint/12045/>

Document Version:

Article (Published Version)

Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

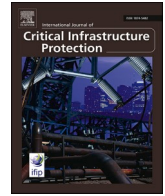
© 2025 The Authors

The aim of the Leeds Beckett Repository is to provide open access to our research, as required by funder policies and permitted by publishers and copyright law.


The Leeds Beckett repository holds a wide range of publications, each of which has been checked for copyright and the relevant embargo period has been applied by the Research Services team.

We operate on a standard take-down policy. If you are the author or publisher of an output and you would like it removed from the repository, please [contact us](#) and we will investigate on a case-by-case basis.

Each thesis in the repository has been cleared where necessary by the author for third party copyright. If you would like a thesis to be removed from the repository or believe there is an issue with copyright, please contact us on openaccess@leedsbeckett.ac.uk and we will investigate on a case-by-case basis.



SPARK and SAD: Leading-edge deep learning frameworks for robust and effective intrusion detection in SCADA systems

Raghuram Bhukya^a, Syed Abdul Moeed^a, Anusha Medavaka^b, Alaa O. Khadidos^{c,d},
Adil O. Khadidos^e, Shitharth Selvarajan^{f,g,h,*} 

^a Department of Computer Science and Engineering, Kakatiya Institute of Technology and Science, Warangal, TS 506015, India

^b Software Engineering, JP Morgan Chase, USA

^c Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

^d Center of Research Excellence in Artificial Intelligence and Data Science, King Abdulaziz University, Jeddah, Saudi Arabia

^e Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah, Saudi Arabia

^f School of Built Environment, Engineering and Computing, Leeds Beckett University, LS1 3HE Leeds, UK

^g Department of Computer Science and Engineering, Chennai Institute of Technology, Chennai, India

^h Centre for Research Impact & Outcome, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, 140401, India

ARTICLE INFO

Keywords:

Supervisory Control and Data Acquisition (SCADA) systems
Intrusion detection system (IDS)
Security
Deep learning
Classification
Optimization

ABSTRACT

Considering SCADA systems operate and manage critical infrastructure and industrial processes, the need for robust intrusion detection systems-IDSs cannot be overemphasized. The complexity of these systems, added to their increased exposure to more sophisticated cyber-attacks, creates significant challenges for continuous, secure operations. Traditional approaches to intrusion detection usually fail to cope, scale, or be as accurate as is necessary when dealing with the modern, multi-faceted problem of an attack vector against SCADA networks and IIoT environments. Past works have generally proposed the use of different machine learning and deep learning anomaly detection strategies to find possible intrusions. While these methods have, in fact, been promising, their effects are not without their own set of problems, including high false positives, poor generalization to new types of attacks, and performance inefficiencies in large-scale data environments. In this work, against this background, two novel IDS models are put forward: SPARK (Scalable Predictive Anomaly Response Kernel) and SAD (Scented Alpine Descent), to further improve the security landscape in SCADA systems. SPARK enables an ensemble-based deep learning framework combining strategic feature extraction with adaptive learning mechanisms for volume data processing at high accuracy and efficiency. This architecture has stringent anomaly detection through a multi-layered deep network adapting to ever-evolving contexts in operational environments, allowing for low latency and high precision in the detections. The SAD model works in concert with SPARK by adopting a synergistic approach that embeds deep learning into anomaly scoring algorithms, enabled to detect subtle attack patterns and further reduce false-positive rates.

1. Introduction

Supervisory Control and Data Acquisition (SCADA) system is basically the backbone of modern industrial control systems. That means, for all practical purposes, they are designed to operate real-time monitoring, control, and automation for variable processes in different vital industries within society, including power generation, water treatment, and oil and gas production [1,2]. These systems play a critical role in maintaining industrial infrastructures and can enable appropriate actions to be taken by operators due to system anomalies. They ensure

production stability, as well as operational safety of facilities through remote maintenance. SCADA systems have been on a spree of high growth in recent times driven by connectivity improvement, data analytics, and technologies that support remote access. More concretely, these improvements have meant more efficient, centralized operations management of distributed and often geographically dispersed industrial systems, greatly boosting productivity while enhancing operational visibility [3,4]. It is at this moment, however, that such capability dispersion also opens up new challenges-particularly in the domain of cybersecurity. Indeed, many of the same features that enable this better

* Corresponding author.

E-mail addresses: akhadidos@kau.edu.sa (A.O. Khadidos), aokhadidos@kau.edu.sa (A.O. Khadidos), s.selvarajan@leedsbeckett.ac.uk (S. Selvarajan).

<https://doi.org/10.1016/j.ijcip.2025.100759>

Received 24 November 2024; Received in revised form 7 February 2025; Accepted 23 March 2025

Available online 30 March 2025

1874-5482/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

management-such as increased remote access and network connectivity-are those exposing SCADA systems to cyber-attacks. Integration of Internet of Things (IoT) into SCADA systems has, however, introduced a new layer of complication regarding these issues of cybersecurity [5,6]. Conventionally, SCADA systems were designed to operate in an isolated environment with very limited or no connectivity to other outside-the-gate networks. This made them quite minimal, by default, in terms of an attack surface. Even today, a strong desire to further operational capabilities and provide insight from the data pushes these systems toward increased connectivity with other IoT devices and networks. This transition comes with benefits like predictive maintenance and optimized resource management at the cost of new challenges that traditional SCADA systems were never designed to operate within. IoT brings in a variety of issues, starting from simple heterogeneity in the field devices connected-each having its own different vulnerabilities-to the lack of uniform security standards that can be enforced across all types of devices and systems [7,8]. Most IoT devices are designed to offer functionality with an affordable cost, which, in turn, means that security features are generally not robust. This form of vulnerability within IoT devices introduces entry points that cyber attackers can leverage, therefore making SCADA system protection far more challenging [9,10].

Also, the criticality of real-time operations of SCADA becomes yet another constraint in a way that security measures should be implemented without process disruptions. The demand for uninterrupted operation may potentially narrow down the range of traditional security best practices, such as regular patching and updating of the software, which can affect systems for some period of time. Intrinsic complexity and interconnectedness, together with the real-time requirements of modern SCADA systems, challenge the deployment of comprehensive security measures [11,12]. Thereafter, this very interesting, urgent research direction appeared, connected with enhancing SCADA system security to develop new solutions capable of mitigating risks without compromising operational integrity. A number of researchers and industrials are working on advanced symmetric/asymmetric encryption techniques, intrusion detection systems, and machine learning models able to identify real-time anomalies among others. In general, the aim is to make the SCADA framework resilient by resisting evolving cyber threats variably.

The neural networks constitute the basic units of deep learning and take their inspiration from the inner structure of the human brain [13]. As a matter of fact, this field of studies has turned out to be very important in modern computational methods. It is really good at recognizing complicated patterns and solving problems that contain high-dimensional data. It means that it defines the architecture of a neural network, the interconnectivity of nodes in layers performing a certain computation. These could be extremely helpful in cases where the job analyses are comprehensive, detection of anomalies in SCADA systems is necessary, for example, whereby other methods often fail to identify a subtle, emerging threat or pattern pointing towards abnormal behavior. This is illustrative of the strategic shift in the deep-learning security paradigm of the SCADA system through aggressive leveraging of the most top-level computational models toward proactive adaptation and response against dynamic threats. Equipped with the capability to process vast data amounts and learn complicated dependencies, they become appropriate to the SCADA environment, where large flows of data from sensors, controllers, and actuators are a common feature. Huge volumes and varieties are the characteristics of data streams generated from such environments; therefore, only against a volume of information will meaningful anomalies be detectable via a strong analytical approach [14–16]. Neural networks learn directly from data, including performance enhanced by experience, and are great at distinguishing between normal operational behavior and potential threats in complex, noisy data conditions.

Detection of cyber-attacks in SCADA IoT is very important from the security point of view to guarantee safety and operational reliability in

highly critical infrastructures for life in modern society. The integration of IoT into the SCADA system provides huge benefits with regards to remote monitoring and real-time acquisition of data. On the other hand, they also raise critical infrastructure vulnerabilities due to more connectivity and various points of entrance with possible vulnerable targets. The solution of this challenge is truly out of the box and hardly achieved using common security [17]. Probably, self-similarity analysis in combination with ML and DL techniques stands a chance to make this work. It is the statistical procedure for the fractal properties, otherwise said, the self-similarity properties in different scales or time intervals of a system or a network. It does have great potential in the identification of deviations which might signal abnormal or malicious behavior within a SCADA IoT network. That is to say, it means leveraging inherent repetitive patterns that are natural in normal network traffic, detecting subtle changes in that pattern, which could indicate a possible cyber intrusion.

Combining ML and DL algorithms with the self-similarity analysis itself creates a hybrid approach that can enormously beef up the detection of cyberattacks [18]. This self-similarity analysis may be used as a preliminary filter, reducing data to those conditions that are possibly anomalous or strange patterns needing close scrutiny. Once these patterns are identified, the ML/DL algorithms would go through the analysis and classify the anomalies with much greater certainty. In like manner, a number of machine learning models will be trained on historical data with a view to learning known attack signatures or behavior patterns, including decision trees, support vector machines, and ensemble techniques [19,20]. On the other hand, deep learning methods-especially complex neural networks like CNNs or RNNs-do a pretty good job of learning difficult nonlinear relationships from data. These deep models would thus be able to extract features from network traffic data that may well elude more simple analysis and enhance total detection capabilities for advanced persistent threats and zero-day attacks.

However, to address the complex threat scenarios that exist in SCADA systems and with emergence of IoT devices, the strategy must combine several approaches. This strategy should be done by strengthening the security of a single IoT device because more often than not, most IoT devices have poor security features. Having an identity that creates equivalent security parameters that may be implemented throughout all associated devices and systems is crucial in order to eliminate the threats and implement changes to device and network security policies [21]. In addition, complex, manifold, and networked systems of modern SCADA require unification of various devices and protocols into an entity and provide it with adequate protection. This integration must be achieved, free of compatibility problems, while at the same time keeping a very tight check on security to ensure the safe transfer of data. The aim is to build a system that upholds the core principles of cybersecurity: ensuring that protecting data from unauthorized access, ensuring accuracy and completeness of data, and ensuring that people have access to system and services.

The kind of security that can be obtained to achieve the objectives of having a single system protection is the multi-layered security model that includes encryption, secure authentication, and threat detection mechanism that integrated deep learning. Neural networks that offer real time anomaly detection systems can also have an added role in this layered approach by having another level of depth which can identify new and emerging threats [22,23]. Surveillance and threat identification are also improved by such systems since it analyses all the data that is continuously received in the SCADA environment to provide alerts whenever activities that are most probable to hinder normal functioning and/ or represent a breach are identified. Additionally, the capability of the neural networks to learn and make new responses to the new data set as time elapses enhances proactive threat management, which makes organizations to be in a position to respond effectively to the ever emerging threats in the cyberspace. Therefore, an optimal solution that incorporates best practices in the DL, secure IoT, and implement

standard security practices to make SCADA better placed to handle new emerging threat models without compromising efficiency and reliability is crucial.

Despite potential and developments that these approaches hold, the intrinsic limitations and related challenges should be solved in this respect for their full exploitation in practice. One of the main challenges when using self-similarity analysis and ML/DL algorithms in SCADA IoT networks is heterogeneity and scale from the data produced by such systems. SCADA networks involve various devices and protocols in IoT, all with different frequencies and levels of reliability across the nature of data generated [24]. These inconsistencies in the nature of data bear the potential for serious complications in developing models that will be generalizable and accurate across diverse network conditions. The high throughput and velocity of the flow of data in SCADA systems require that detection models work with minimal latency to avoid delays in responses that might compromise the integrity of the system.

Although improved accuracy has been observed in deep learning detection algorithms, the computational cost is very high, often unaffordable in resource-limited environments. Therefore, there is a need either to develop more computationally efficient models or use hardware accelerators. The frequent retraining of models, however, brings in operational complexity and skilled people who might not be affordable for all organizations, especially those that operate in resource-constrained settings. On the whole, while the self-similarity analysis combined with the ML/DL techniques is one of the best approaches for detecting cyber-attacks within SCADA IoT networks, turning it into practice calls for continued research and development by overcoming the challenges above for scalable real-time protection against the unique demands placed upon the SCADA environments. This research is motivated by the realization that the current security frameworks are somewhat limited in their capability to handle the complexity and dynamic nature of the cyber threats that increasingly target SCADA environments [25]. Current methodologies have tended to rely heavily on signature-based detection methods, which already have knowledge about the patterns of an attack and are therefore incapable of discovering zero-day exploits and novel intrusion strategies. The deeper capabilities of deep learning in terms of the ability to learn from vast amounts of data, recognize intricate patterns, and propose a more proactive approach to cybersecurity are harnessed in this research work [26]. Then again, optimization techniques can be applied that improve the training of the model to learn from a few data, which typically is the problem in cybersecurity, where a complete dataset of an attack cannot usually be found. More importantly, this capability is critically needed in SCADA systems [27,28], where the operational data is often limited and highly contextual.

Major Contributions:

- This paper introduces two new models of intrusion detection, SPARK and SAD are meant to enhance, through ensemble-based deep learning and anomaly scoring mechanisms, the security in SCADA and IIoT environments against ever-evolving complex cyber threats.
- The SPARK model adopts adaptive learning fused with strategic feature extraction for satisfying the requirements of scalable and efficient anomaly detection of large-scale industrial datasets. SAD complements SPARK with sophisticated anomaly scoring that targets complex attack patterns for reduced false positives.
- By using the SWaT and WUSTL-IIoT datasets, this paper showing better detection performance compared to that of traditional and contemporary IDS solutions. The same hyperparameter tuning, evaluation metrics, and statistical validation techniques are followed strictly in this study to ensure a fair comparison.
- The research also seeks to enhance the resiliency of the SCADA systems against cyber threats, both in the present and emerging contexts, by developing a better understanding of how advanced machine learning approaches can be effectively used in real-time operational environments.

- Results from this study can, therefore, shape policy and best practice for the industry, providing stakeholders with new tools and strategies that assist in protecting their systems and maintaining continuity of operations against determined cyber threats.

Outline:

The organization of the paper is such that the pressing issue of intrusion detection in SCADA systems has been viewed through advanced machine learning and deep learning approaches in a very structured manner. Section 2 will review the related literature concerning the existing machine learning and deep learning models applied to intrusion detection in the SCADA environment by describing their methodologies along with their respective strengths and limitations. This review will set a baseline understanding of the current status of research in this domain and thus provide the platform for the proposed contribution of the study. Section 3 will discuss the proposed methodology, covering the development of a hybridized deep learning model integrated with optimization techniques designed especially for enhancing the security of SCADA systems against cyber threats. The efficiency of the proposed model will be proved in this section using different metrics, and its superiority to the existing approaches by presenting the detection rates, false positives, and computational efficiency. Finally, Section 5 concludes the research findings on implications of results and future work toward enhancing intrusion detection in SCADA systems, pointing out that innovation concerning the addressed field is a never-ending process.

2. Related works

Other recent works also pointed out the inclusion of hybrid approaches which integrate the best of machine learning with deep learning methods. Besides, integration of optimization techniques within such models remains a key area of concentration for model parameter tuning to perform better with minimum computational overhead—a critical issue in real-time SCADA applications [29,30]. In a nutshell, this stream of research in this area underlines continuous innovation and adaptation of intrusion detection techniques in a dynamically changing threat landscape that SCADA networks have to face. This section outlines the specifics of these state-of-the-art approaches, their methodologies, relative benefits, and challenges—thus laying the ground for the proposal at hand, which would like to make another contribution to this very critical area of cybersecurity.

Ragab, et al. [31] propose the NGCAD-EDLM approach, where the acronym refers to Next-Generation Cybersecurity Attack Detection by means of an Ensemble Deep Learning Model; this is aimed at solving newly arising safety vulnerabilities because of integrated legacy ICT systems within industrial IIoT contexts. The study presents that the development and deployment of new safety technologies in power control systems are likely to face serious cyberattacks; therefore, it has become urgent to establish an efficient cybersecurity system in place. As the critical role of these systems in ensuring reliable power delivery is so high, the stability and efficiency of the proposed cybersecurity model is of prime importance. The NGCAD-EDLM technique has been devised in order to support the automated detection of cyberattacks, so as to improve the security posture of IIoT infrastructures. It follows the use of a principal approach for primary data normalization through min-max normalization, which is an important step since it scales the data for appropriate analysis. The synergy between these two methods within the framework of an ensemble is expected to yield improved performances both concerning detection rates and concerning the reduction of false positives.

Abdulganiyu, et al. [32] present the XIDINTFL-VAE framework that comprises CWFL and VAE, embedding XGBoost to tackle challenges imposed by class imbalance in intrusion detection systems. This proposal is very likely to further improve the capability of detection for minority class intrusions without sacrificing overall robust performance

across classes. Unfortunately, most of these approaches tend to face significant challenges in balancing precision and recall on highly imbalanced datasets. As a result, these strategies may lead to a high rate of false alarms or missed cases. This may pose serious problems in security applications where the cost of errors could be extremely high. The XIDINTFL-VAE framework directly addresses this important gap by synthesizing data from the most difficult cases in the minority class. This tailor-made approach will enable a more subtle analysis of the subtlety related to the rare intrusions, which will be useful for improving the detection rate and overall classifier performance. Huang, et al. [33] have proposed a study wherein, for the very first time, there is the use of sequence feature construction algorithms to represent explicitly the information of sequence features, laying a foundation for a truly effective intrusion detection system.

The proposed LSTM networks and feed-forward neural networks work in tandem so as to retain critical sequence information while the dimensions in the output dynamically adapt. This network architecture of two kinds works in an effective mapping of processed information to classification labels. Finally, the simulation comparison results show that the designed IDS has much higher packet capture rate per second when compared to three other existing systems. Proposed system achieves an impressive packet capture rate of 7000 packets per second with intrusion rates of 10 and 22 % and maintains the system occupancy rate at 23 %. This performance metric value is quite impressive and stands in support of the efficiency and effectiveness of the proposed approach in handling intrusion detection tasks with different attack intensities. The results of this study indicate that the proposed mechanism performs significantly better than the competition in intrusion detection and response. From this, the authors provide several effective solutions to the crucial challenges of intrusion detection and response in a timely manner, utilizing an appropriate combination of latest neural network architectures and sequence features extraction. The obtained results show the practicality of the proposed system for use in the real world; at the same time, they focus on its potential as a strong countermeasure that can help enhance the security posture of networks against an increasingly sophisticated landscape of cyber threats.

Mesadieu [34] propose a DRL framework for anomaly detection in SCADA networks. By leveraging a "Q-network," it positions itself at an advantage to achieve state-of-the-art performance in recognizing patterns from complex tasks, very essential for effective anomaly detection in industrial control systems. Hence, the integration with DRL adds the capability for continuous improvement in the detection capability through interaction with the environment and thus adaptation of the model to evolving threat landscapes and operational conditions. These authors have conducted experiments on two publicly available datasets for the validation of their proposed solution. Such validation is quite important since it proves that this model will be able to generalize on other types of data in a real-world scenario. That is a very good contribution to cybersecurity in SCADA systems through deep reinforcement learning by enhancing the anomaly detection mechanisms. The DRL framework proposed in this paper has focused on state-of-the-art performance in pattern recognition and is expected to help enhance the resilience of the SCADA network against possible cyber threats and, in turn, contribute to security and reliability in critical infrastructures.

Zaman, et al. [35] provides an overall design framework for the ML-based IDS for SCADA-based power systems. Fully aware of the possibilities opened up by ML techniques for enhancing security measures, the authors underline some of the intrinsic limits in the development of ML models, mainly related to the need for customized methodologies in data preprocessing and training. These challenges are addressed appropriately by the proposed framework, which embodies a few key aspects of modeling. The authors have performed various experiments to validate their proposed design framework using a publicly available dataset from ORNL specifically related to SCADA-based power systems. Empirical validation provides the opportunity to present the

effectiveness of the proposed design framework in a real-world environment, and at the same time, it acts as a benchmark for not only all proposed design frameworks but also several existing IDS solutions to compare the performance.

Sangoleye, et al. [36] have determined the limitations of the current approaches using Machine Learning, especially in intrusion detection within ICS networks. They determine that most of the traditional methods require very frequent manual retraining and hardly keep pace with the dynamic nature of evolving cyber threats. In the paper, a new research has been carried out on the applications of a variety of models such as Deep Q-Network, Double Deep Q-Network, Dueling Double Deep Q-Network, REINFORCE, Advantage Actor-Critic, Proximal Policy Optimization. Each of these models reflects a different approach to the exploitation of reinforcement learning principles in network intrusion detection and provides extensive study of the DRL possibilities for enhancing the security of ICS. In such a way, the authors try to reveal the strengths and weaknesses of each model with regards to practical applications and provide valuable insights about their efficiency in intrusion detection tasks. Suffice it to say, the work of these authors marks a significant step in applying advanced DRL techniques within ICS security frameworks. Emphasizing both autonomous learning and adaptability, their research has secured intrusion detection systems' responsiveness against the rapid pace of threat evolution-much needed for hardened critical infrastructure against cyber-attacks. This research also builds on the literature in terms of the current state of DRL model applications in cybersecurity applications and sets up a foundation for follow-on research work in terms of further optimizing such models for practical use.

Ali, et al. [37] pointed out an important gap in the present research landscape, where much attention has gone to binary classification problems while multi-class classification remains a challenging and actively evolving area. The proposed instance-based intrusion detection technique tailored for IDS-ICS is especially for SCADA networks to meet the challenge. It is named ICS-IDS, for overcoming multi-class imbalanced classification challenges. Most important identification of various intrusion types in ICS environments depends on this approach. The proposed technique of ICS-IDS has two major portions: preparation of data and detection. The portion for preparing the data applies several advanced techniques to enhance the quality and further analyzability of the dataset. Normalization first scales the data to equal importance of all features during model training. Dimensionality reduction is then used to avoid the problem of the curse of dimensionality by retaining only the most informative features with the intent of improving model performance. It then implements the methodology for k-nearest neighbors and resamples the dataset to balance out properties and include more representation of minority classes.

Yalcin, et al. [38] have underlined that cyberattacks against network-based communication structures form a critical vulnerability for industrial equipment and operations in ICS. They also pointed out that this type of attack may further cause massive disruption or even sabotage to manufacturing processes; hence, they ring an alarming call for enhanced security to be executed on these systems. With the continued digitization of ICS, the exposure to malicious actors keeps on increasing, and, thus, strong security solutions have been called for in ensuring safety, especially through Intrusion Detection Systems. Given that sophistication in cyberattacks keeps on evolving, the authors assert that industrial companies are obliged to innovate through the adoption of sophisticated solutions such as attack detection systems using artificial intelligence. Adaptability proves to be an important element because critical industrial operations are among the favorite targets of cybercriminals. Thus, the current research aims at developing an AI-based IDS that has been efficient in enhancing security in SCADA systems while ensuring high accuracy in the threat detection. The authors are striving to benefit from all AI technologies in developing a solution that is not only effective in intrusion detection but also learns with time, for new and emerging threats, hence contributing to the

resilience of industrial operations. This also represents an important research contribution, as it addresses urgent security needs and tries to push the state-of-the-art in methodologies for intrusion detection in critical infrastructure environments. The manuscript thus contributed to the ongoing discussion regarding the need to embed AI into cybersecurity frameworks, with the ultimate goal of raising protective means for industrial actors.

Islam, et al. [3] have proposed a new approach for intrusion detection in complicated data environments through the development of an efficient Long Short-Term Memory-based Sparse Variational Autoencoder technique, LSTM-SVAE. A new technique is developed, which could effectively extract relevant features from intricate data patterns and solve problems normally brought forward by high-dimensional and noisy datasets. Hence, by leveraging the strengths of LSTM networks in finding subtle and complex relationships within the data, this is a critical aspect that the authors achieve for effective intrusion detection. Following feature extraction, the authors design a Bidirectional Recurrent Neural Network with Hierarchical Attention-sequence Intrusion Detection. Such an architecture is meant to proficiently find potential intrusions by incorporating advanced memory capabilities together with mechanisms of focus. The hierarchical attention mechanism further helps the model in performance by allowing it to prioritize the features pertinent for intrusion detection analysis. Next, the authors introduce a Cognitive Enhancement for Contextual Intrusion Awareness module, CE-CIA, which refines the initial predictions from the BiRNN—HAID. The proposed component adopts cognitive principles to balance sensitivity and specificity when accomplishing intrusion detection, hence being able to keep false alerts as low as possible.

By improving the dependability of the detection system, the overall effectiveness of the intrusion detection framework in underlining real threats while minimizing unnecessary alerts thus gets enhanced by means of CE-CIA. Wali, et al. [39] provide a broad survey of prevalent cybersecurity vulnerabilities and attacks that cloud-based SCADA systems face, furthering the needed knowledge in understanding the security landscape for these infrastructures. The study points to four important factors of vulnerabilities that make those systems potentially vulnerable: connectivity with the cloud service, shared infrastructure, malicious insiders, and security related to SCADA protocols. Further, the authors classify cyberattacks that target these systems into five significant groups, namely hardware attacks, software attacks, attacks focusing on communications and protocols, control process attacks, and insider attacks. This can be likened to a detailed categorization of attacks, which forms the very basis on which focused security strategies are developed. In addition, the authors of the paper go ahead to identify the various types of attacks and propose various security solutions aimed at mitigating the impact of cyber-attacks against cloud-based SCADA systems. To this end, by proposing security solutions, the authors have given practical suggestions that could go a long way in reinforcing the resilience of the SCADA systems against such emerging threats.

Sogut, et al. [40] present a very interesting study physically emulating-real water plants on a smaller scale—a carefully developed Testbed environment with a SCADA system. This novelty allows going deeper into the exploration of vulnerabilities and resilience of such systems in controlled conditions, giving insights into the dynamics of their operation under conditions of simulated attack. To realize this aim, the authors designed five different attack scenarios, each using a different Distributed Denial of Service attack: TCP, UDP, SYN, IP spoofing, and ICMP flooding. Since the SCADA system was intentionally made to malfunction with those specific attack vectors, this research shows well the risk that these systems can be exposed to in real-life conditions. The approach focused not only on underlining different methods of DDoS but also supported capturing subtleties regarding their impact on the operation of SCADA. Besides the attack scenarios, the authors considered a baseline scenario that depicts normal behavior of the SCADA system while considering no interference. This serves as a baseline for the authors, so that they are able to compare how much each

attack scenario disrupts the operation of the system in question. While these attack scenarios were being performed, the network of the SCADA system was closely monitored, and the network data was collected and recorded. Indeed, this intensive collection of data forms a firm basis on which to consider the impact each of the different DDoS attacks has on the performance of the SCADA system and therefore, by virtue of that, provides valued contribution to the existing literature on the security and resilience of SCADA systems against cyber threats.

Sahani, et al. [41] conduct a broad review on ML-based IDS for smart grids. They discuss several key aspects with care that form part of the necessary understanding of the role of ML in enhancing security for these complex systems. First, ML-based IDS applied to the transmission and distribution sides of power components in smart grids is explored, while strongly focusing on the identification and addressing of inherent security vulnerabilities. This attention to data illustrates that there is, in fact, a need for quality input if, indeed, ML techniques are to be used successfully to guarantee that the models identify and respond correctly to a potential intrusion. Additionally, the survey goes into detail on the wide breadth of ML-based IDSs that have been deployed by the surveyed literature, from a superficial look at the various algorithms and methodologies employed by researchers on the subject to detailed discussion. The wide representation not only serves to present the flexibility of approaches toward addressing security challenges for smart grids using ML but also as a useful resource for practitioners with interests in implementing effective detection systems.

While there has been significant development regarding IDS for SCADA networks, much time finds several research gaps still exist. One major pitfall lies in the shallow investigation of the hybrid model that combines several ML and DL techniques to solve the peculiar challenges of the SCADA system. The literature, on one hand, focuses mostly on traditional machine learning approaches or, on the other hand, deep learning methods separately and does not give much importance to benefits that could be realized with a combination of these paradigms. This indeed is an indication of how broad a framework is needed that could capitalize on both methodologies for enhanced detection rates and minimal false positives in different operational scenarios. Also, most of the current research designs lack a deep understanding of the cyber threats to SCADA systems as being dynamic and evolving in nature. This is because, although several of these studies adopt static datasets for training and evaluation, it does not serve the realistic purpose underlying the detection and response to new attack vectors in real time.

Indeed, methodologies that will make real-time learning and adaptation possible are definitely those that future research will need to focus on to respond to the demand of changing systems, which evolve in the context of emerging threats. Another important shortcoming is sufficient attention and emphasis on interpretability and explainability of IDSs in the SCADA environment. Most of the current techniques are targeted at performance evaluation metrics like accuracy and recall; however, their decision-making process remains mostly in a nontransparent manner. This points to one of the most important gaps that need to be filled since such understanding of the reasons and ways certain decisions are taken by stakeholders is so crucial for improved risk management and response strategies.

The related work in the area of intrusion detection for SCADA systems and IIoT environments has pros and cons, based on the methods and technologies used. One of the most obvious benefits of using machine learning-based IDS is that it can potentially be applied to detecting unknown threats by defining anomaly behavior patterns. Traditional approaches, therefore, heavily rely on predefined attack signatures that is, they can recognize only known threats. Such capability of handling and analyzing big data is particularly useful during scenarios when SCADA systems are generating huge streams of real-time data, including sensor readings, control commands, and other system logs. Furthermore, deep learning methods, especially CNNs and RNNs, have shown outstanding performances in the detection of complex attack scenarios with sequential or spatial patterns in data. Those methods achieve

significant performance in dealing with cases of SCADA networks, where long-term dependencies and interdependencies between devices and processes are likely to prevail.

Another, though related, area where those methods are also very prominent is machine learning-based IDS in SCADA and IIoT systems. Among those challenges, one of the most important is that it is computationally costly to train and deploy a machine learning model, especially a deep learning model. High computational demand may cause the latency issues in real-time detection, which is hard to meet the stringent time requirements of the SCADA systems controlling the critical infrastructure. It also makes it very hard to debug or fine-tune the models in real-world operational environments due to the opaqueness in decision processes. Labeled data for infrequent attack scenarios may be very problematic to have in practice. False positives, although usually not very important in non-industrial settings, are quite troublesome in SCADA systems. The false alert may initiate unnecessary shutdowns or corrective actions with possible impacts on system operations or even damage to critical infrastructure.

3. Proposed spark methodology

In this modern world, where computation in neuroscience is constantly improving, it is a matter of utmost urgency and importance to further elaborate an efficient and effective model for spike encoding threshold computation in order to advance our knowledge with respect to the neural process and the way information is depicted in the brain. This proposal presents new work concerning a hybrid model, one which will be built from the power of two state-of-the-art methods: the Spike Encoding Adaptive Regulation Kernel (SPARK) and the Scented Alpine Descent algorithm. The main novelty of the present study is a synergistic combination of the two approaches, which enhances not only the precision of spike threshold determination but also the adaptability and robustness of the model in different dynamic conditions. A blend of strengths like this from SPARK and SAD can hopefully mitigate some of the weaknesses exhibited by each of these methods to offer a better real-world solution in neural data analysis and signal processing, among other related areas.

The SPARK model introduces a novelty for spike encoding regarding adaptive thresholding with regard to characteristics of input signals. It uses dynamic adjustments in encoding parameters using advanced statistical methods, given that neural inputs are of a varying nature. This will make the model more sensitive to fluctuations in stimulus intensity and frequency for better encoding of information. The key strength of SPARK is that it can combine the temporal and spatial features of neural activity. In particular, it is ideal for complex datasets where the other traditional methods with fixed thresholds cannot detect critical variations. SPARK forms one threshold determination process that is much more sensitive to context; thus, it greatly enhances the fidelity of neural encoding—a solid foundation for the proposed hybrid model.

Complementing SPARK, the contribution of SAD brings a unique optimization approach conceptually rooted on principles of olfactory navigation and Levy flight mechanics. This method allows for an efficient exploration of the solution space so that the spike encoding thresholds will not only be optimal but also resilient to local minima pitfalls commonly encountered in traditional optimization techniques. The novelty of using odor-based cues in guiding the search process across a complex space brings fresh dimensions into the processes of optimization and furthers a more holistic understanding of the linkage between neural spikes and environmental stimuli. SPARK thereby gives an adaptive regulation in a complementary way to the strong optimization due to SAD. Hence, SPARK can form a powerful framework to model fine intricacies of spike encoding threshold computation. This hybrid model is novel because it is integrative and hence effectively marries the adaptiveness of SPARK with the advantages of exploration by SAD. This kind of fusion alone has amplified the individual strengths of each technique to create new capabilities which enhance

performances. For instance, the fact that this model automatically adjusts spike thresholds in a dynamic way, while simultaneously optimizing the search for those thresholds, represents an important extension beyond methodologies that, within other approaches, occur in isolation or rely upon a static parameter. Furthermore, importing mechanisms inspired by biology into the computational framework introduces even another layer of sophistication, making the model's alignment with natural neural processes even closer. This work, therefore, constitutes one of the most important and significant steps in pursuit of the most authentic and efficient spike encoding methods that would clear the path for innovative applications both in research and practical settings regarding neuroscience and beyond.

SPARK is particularly appropriate for the study in hand because of its singular design focused on SCADA system-related and IoT-driven ambient challenges. One of the most important features of SPARK is scalability with the size and complexity of the data. The SCADA systems, especially those integrated with IoT devices, generate huge amounts of data from diverse sensors and controllers. SPARK can efficiently process this data using adaptive learning mechanisms, which hierarchically extract the most important features of the data, so it would not overwhelm the system with the anomaly detection process. This is in contrast to many IDS models, which either employ fixed feature sets or need manual tuning for adaptation to the varying levels of complexity in the data. SPARK is flexible and, hence, might change according to the operational environment; thus, it results to be very suitable for real-time dynamic industrial systems, where network configurations, device behavior, and attack strategies might change over time. Unlike in most conventional models, which may need retraining or even hand intervention to get used to new patterns of data, SPARK assures that its dynamic learning abilities make it effective and responsive with little performance degradation.

Most of the existing IDS models, especially those deep learning-based ones, are easily burdened with a high computational overhead that slows down the detection process especially in an IoT environment where devices usually possess weak processing power. Being lightweight in feature extraction and efficient in learning mechanisms, SPARK can be executed on IoT devices or at edge nodes without inducing any extra significant delay. Besides, SPARK also showed great strides in handling false positives, which are very common with most IDS solutions. Most of the traditional IDS, due to their simplicity and mostly rule-based or signature-based approaches, lack both adaptability and precision; thus, they often tend to produce either a large number of false positives or detection failures both of which are quite disastrous in the case of industrial control systems (Fig. 1).

3.1. Spike wasserstein adversarial robust cognition (SPARK) classifier for intrusion detection in SCADA systems

The SPARK Classifier is a new intrusion detection methodology for SCADA systems. This technology effectively embeds the power of spike-based neural networks with the principles of WGAN, including its most salient features, for constructing a classifier that is both robust and efficient at identifying complicated patterns related to potential intrusions within SCADA systems. The SPARK Classifier exploits the peculiar properties of spiking neurons, which are much closer to their biological counterparts, enabling temporal processing, among others, for higher energy efficiency. This is especially a bonus in real-time monitoring scenarios typical of SCADA systems, where timely responses are crucial for security threats. One of the major contributions of SPARK Classifiers is the novel use of adversarial training for enhancing intrusion detection robustness.

Generative Adversarial Networks (GANs) have been explored in the last few years and show great promise to improve many applications in machine learning, mainly within the area of anomaly detection and IDS. Application of GANs in IDS models is justified, since they can generate realistic synthetic data that can enhance model robustness by learning

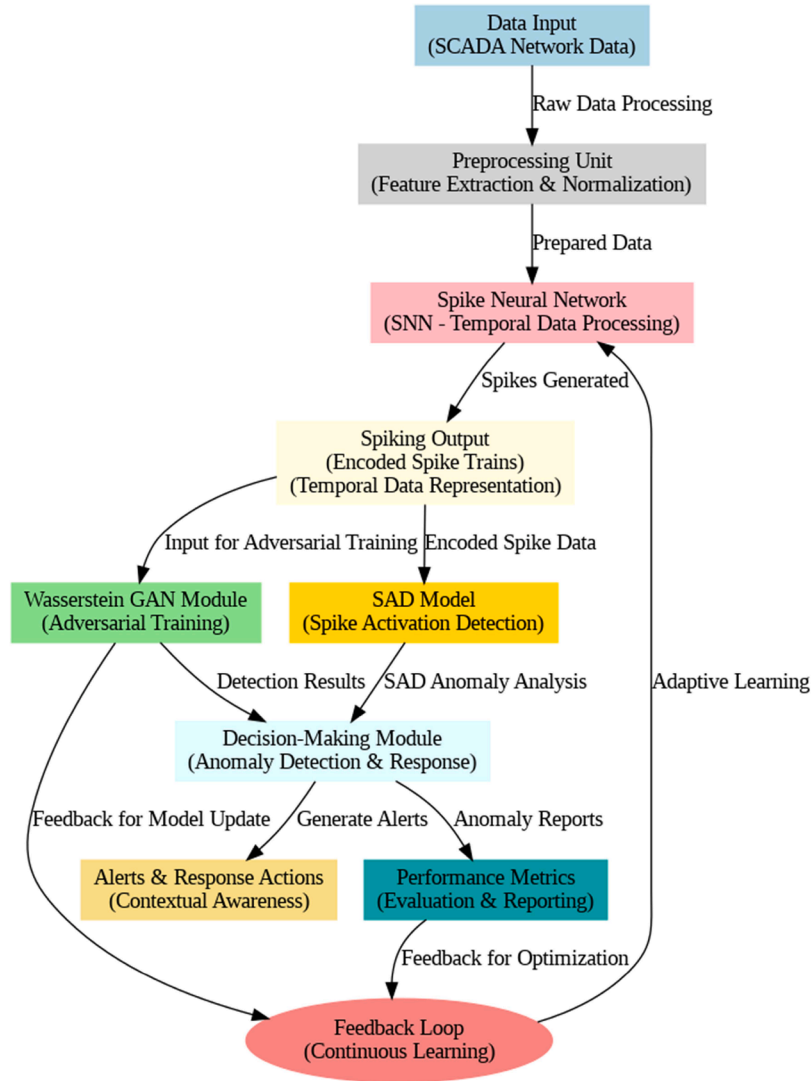


Fig. 1. Overview of the proposed model.

the complex data distribution. This limitation has been one of the motivations to investigate GAN-based methods because GANs are capable of generating realistic synthetic attack samples, enabling IDS models to generalize better when dealing with new and unseen threats. This has the proposed IDS framework integrated with GANs, which is especially important for SCADA systems where cyber threats evolve very fast. It therefore makes the framework highly adaptable and resilient, hence its high efficacy in securing industrial control systems and IIoT environments.

Their decision to consider GANs in the IDS models lies with their remarkable strength in tackling problems of data scarcity and class imbalance, which are pervasive challenges in intrusion detection. Any malicious activities within the SCADA networks occur relatively much fewer in number as compared to the normal operations. The resultant datasets used are usually heavy-skewed toward the benign class. Such imbalances can even make the traditional machine-learning models not quite effective in recognizing such rare but very critical attack patterns. In this way, the adversarial training process in GANs makes the generator learn to create increasingly realistic data, and the discriminator refines its ability to distinguish between real and fake samples.

This model, taking as a measure of divergence between the distributions of normal and malicious activities, overcomes Wasserstein distance for all traditional problems of GANs: mode collapse and instability of training. In particular, the integration of spiking neural networks into

this framework empowers it to enhance its capability for handling time-series data in a way that captures temporal network traffic dynamics. The dual approaches will not only improve the classifier's known threat detection capability but also enhance its generalization for previously unseen attack vectors, a key requirement in the dynamic cyber threat landscape.

The novelty of the SPARK Classifier is in the unusual integration of spike-based processing with adversarial robustness, hence being very different from the existing models. Traditional machine learning and deep learning approaches face many temporal aspects of SCADA data and the ever-changing nature of attack strategies. In contrast, the SPARK Classifier uses spiking neurons to extract time-dependent features, making the classifier more sensitive in detecting subtle anomalies indicative of a breach. Moreover, SPARK follows the adversarial training paradigm whereby there is assurance that a model will be resilient against adversarial attacks, something particularly relevant given sophisticated techniques adopted by intruders. Furthermore, the model is very adaptive; it can learn from emerging threats continuously to adapt, as the threat landscape keeps on dynamically changing in SCADA systems. The benefits of the SPARK Classifier go beyond its improved detection capability. First, its architecture was designed to ensure computational efficiency during deployment in resource-constrained SCADA environments.

The spiking nature of this network allows sparse communication and

low power use, hence successfully meeting the operational requirements of many industrial control systems. The operational framework of SPARK Classifier focuses on effective network traffic data analysis. First, the information fed from SCADA systems, such as control signals, telemetry, and event logs, is preprocessed into features which would tell the model whether or not the behavior in question is normal or anomalous. This time-series data is then fed through the spiking neural network component to produce spikes indicative of the timing and intensity of activities throughout the network. These spikes feed the Wasserstein adversarial training mechanism, wherein the classifier is trained to discriminate against normal traffic versus various attack patterns.

During learning, the model optimizes its parameters with the Wasserstein loss in such a manner that the latter provides a much more informative gradient to the traditional loss functions for model optimization. The SPARK Classifier develops, through this process, a strong decision boundary to separate legitimate activities from malicious activities. Since the classifier performs real-time monitoring, the detected anomalies could trigger immediate alerts and responses, hence enhancing the security posture of the SCADA systems. In a nutshell, SPARK represents the leading revolution in intrusion detection for SCADA systems by using synergistic performance between spiking neural networks and Wasserstein adversarial training. The architecture of the system overcomes not only some of the limitations inherent in previous models but also provides unparalleled instruments for protecting this vital infrastructure from the ever-evolving threat of cyberattacks.

SPARK Classifier is representative of state-of-the-art technologies in industrial cybersecurity in that it has strong capabilities related to

detection, adaptability to new attack patterns, and a high degree of efficiency from a computational point of view. As shown in Fig. 2, several working modules define the SPARK architecture in a structured manner to present overall efficacy of intrusion detection in SCADA systems. This model thus takes its root from its SNN, which will process inbound data in the form of biological neurons. In SNNs, information is conveyed by discrete spikes rather than continuous signals, thus efficiently encoding temporal data. This is particularly useful in SCADA, where network traffic and operational signals are typically time-series-bound. The raw data streams of control signals, telemetry readings, and event logs are firstly prepared into a format that can be represented by spiking neurons by the SNN module. This module uses various mechanisms, such as temporal coding and spike-timing-dependent plasticity that allow it to adaptively learn from input data over time. The SNN will provide spikes as output while processing the incoming spikes, indicative of its current state and therefore give a rich representation of normal as well as potentially anomalous behaviors. The outputs from the spiking network feed into the WGAN module that forms the backbone of the adversarial training framework of the SPARK Classifier.

In the WGAN module, the classifier plays a game with a two-player generator and discriminator. With this setting, the generator generates adversarial samples, resembling potential intrusions, by learning features from the output of SNNs. The motivation here is to synthesize a diverse dataset composed of not only legitimate traffic but also different attack patterns. In contrast, the discriminator scores both the genuine data from SCADA systems and the synthetic samples generated by the generator against each other in order to discriminate between normal activities and malicious activities with high accuracy. The Wasserstein loss function provides more stability within this adversarial setup and

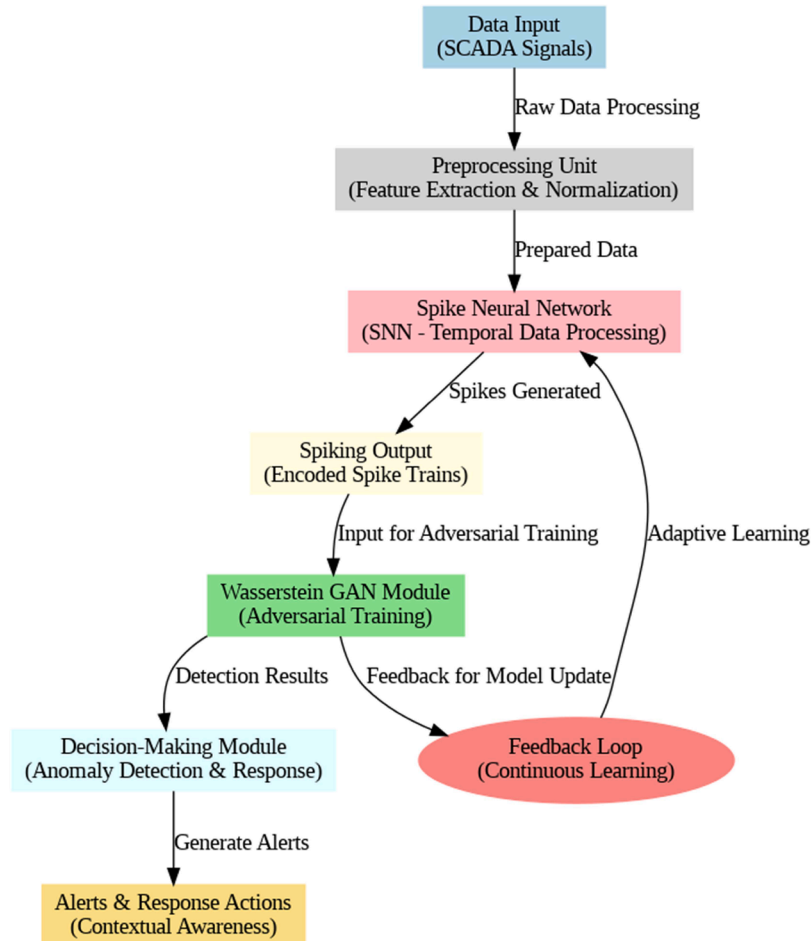


Fig. 2. Flow of the proposed SPARK model.

thus leads to an effective training mechanism that allows the model to learn meaningful representations for both normal and anomalous behavior, not suffering from mode collapse problems so often present within standard GANs frameworks.

A feedback loop in SPARK allows the Classifier to learn continuously to achieve strong performance by updating model parameters with new data, given performance related to a discriminator that recognizes which of the samples are real and which are generated. This turns out to be very critical for enhancing the classifier's novel intrusion strategies detection capabilities that may not have been presented during its initial training. This feedback mechanism, integrated with SNN and WGAN modules, in essence enables the SPARK Classifier to evolve with evolving attack patterns in order to maintain a high accuracy level in detection. The preprocessing unit of the SPARK Classifier prepares feature extraction and normalization apart from the main modules, filtering out noise and irrelevant information with the hope of improving the quality of input feeding into the SNN. Advanced preprocessing may then apply dimensionality reduction to focus only on the most relevant salient features that contribute toward intrusion detection so that the model is not overwhelmed by extraneous data. Then, the pre-processed data is fed into the spiking module for temporal processing, and it gets a further transformation to the spike trains with temporal characteristics.

Its output feeds into a decision module incorporating classifier predictions with SCADA contextual information. This latter module should enable ranked alerts to be generated regarding the anomalies detected, their priority, and recommendations on response actions. By integrating the contextual awareness, SPARK Classifier can determine whether an anomaly represents a serious threat and requires intervention, or a non-threatening anomaly due to normal fluctuations of operation. This holistic intrusion detection approach will guarantee that SPARK Classifier identifies not just potential intrusion but allows for timely and valid response to the threat so as to keep SCADA systems secure.

In this classification technique, the input SCADA network dataset is initialized with the feature vectors as mathematically represented in the following equation:

$$\mathbb{S} = \{(\mathbf{p}_i, \mathbf{q}_i)\}_{i=1}^N \quad (1)$$

Where, $\mathbf{p}_i \in \mathbb{S}^d$ indicates the input feature vector and $\mathbf{q}_i \in \{0, 1\}$ represents the output label defines the instance of intrusion. As a consequence of this, the generator function is estimated from the latent space with the data as described in below:

$$\hat{\mathbf{p}} = \mathcal{G}(\mathbf{x}; \mathbb{Z}_{\mathcal{G}}) \quad (2)$$

Where, $\mathbf{x} \sim \mathcal{N}(0, 1)$ and $\mathbb{Z}_{\mathcal{G}}$ denotes the generator parameter. As a consequence of this, the discriminator function is estimated for determining that whether the input data is real or fake based on the following equation:

$$\mathbb{S}(\mathbf{p}; \mathbb{Z}_{\mathcal{G}}) \in [0, 1] \quad (3)$$

Where, $\mathbb{Z}_{\mathcal{G}}$ denotes the discriminator parameter. Moreover, the Wasserstein loss function is estimated as follows:

$$\mathbb{L}_{\mathbb{S}} = \mathbb{E}_{\mathbf{p} \sim \mathcal{P}_{\text{data}}}[\mathbb{S}(\mathbf{p})] - \mathbb{E}_{\hat{\mathbf{p}} \sim \mathcal{P}_{\mathcal{G}}}[\mathbb{S}(\hat{\mathbf{p}})] \quad (4)$$

Where, \mathbb{S} indicates the loss parameter among the real and generated data. Moreover, the gradient penalty factor is computed as follows:

$$\mathbb{L}_{\mathcal{G}} = \mathbb{E}_{\hat{\mathbf{p}} \sim \mathcal{P}_{\mathcal{G}}}(\|\nabla \mathbb{S}(\hat{\mathbf{p}})\|_2 - 1)^2 \quad (5)$$

Where, $\tilde{\mathbf{p}}$ is uniformly sampled among the real and fake data. Consequently, the total discriminator loss function is computed with the gradient penalty factor using the following equation:

$$\mathbb{L}_{\mathbb{S}_{\text{tot}}} = \mathbb{L}_{\mathbb{S}} + \xi \mathbb{L}_{\mathcal{G}} \quad (6)$$

Where, ξ is the balancing factor that is used to balance the gradient penalty. As a consequence of this, the generator loss function is computed that supports to maximize the discriminator output based on the following model:

$$\mathbb{L}_{\mathcal{G}} = -\mathbb{E}_{\hat{\mathbf{p}} \sim \mathcal{P}_{\mathcal{G}}}[\mathbb{S}(\hat{\mathbf{p}})] \quad (7)$$

Where, $\mathbb{L}_{\mathcal{G}}$ indicates the generator loss function that is used to maximize the output of discriminator. In addition to that, the He initialization is performed for updating weight values in the network layers, which is mathematically expressed as follows:

$$\text{AptCommandmathcalw}_1 \sim \mathcal{N}\left(0, \frac{2}{n_i}\right) \quad (8)$$

Where, n_i represents the number of input units in the network layer. Moreover, the spike activation function is estimated with the following model:

$$\delta(\mathbf{k}) = \begin{cases} 1 & \text{if } \vartheta(\mathbf{k}) \geq \vartheta_{\text{kh}} \\ 0 & \text{Otherwise} \end{cases} \quad (9)$$

Where, $\vartheta(\mathbf{k})$ represents the potential membrane, and ϑ_{kh} denotes the spike threshold. Moreover, the loss function for the spiking network is also computed as shown in the following equation:

$$\mathbb{L}_{\text{spi}} = -\sum_{k=1}^K (m_k \log(\delta(\mathbf{k})) + (1 - m_k) \log(1 - \delta(\mathbf{k}))) \quad (10)$$

The combined loss function for the overall SPARK model is estimated using the following equation:

$$\mathbb{L}_{\text{loss}} = \mathbb{L}_{\mathbb{S}_{\text{tot}}} + \mathbb{L}_{\mathcal{G}} + \mathbb{L}_{\text{spi}} \quad (11)$$

The final classified output for intrusion detection in SCADA systems is obtained as represented in the following form:

$$\hat{\mathbf{m}} = \varphi(\text{AptCommandmathcalw}_{\text{out}} \times \delta(\mathbf{k})) \quad (12)$$

Where, $\text{AptCommandmathcalw}_{\text{out}}$ indicates the weight value of the output layer and φ is the activation function.

Scented Alpine Descent (SAD) for Spike Encoding Threshold Computation

The Scented Alpine Descent (SAD) is a special hybrid optimization model of Spike Encoding Threshold Computation for bringing more fine tuning into threshold values pivotal for spike encoding in SNNs. Estimation of spiking thresholds plays an important role in determining how continuous input from SCADA systems will be converted into spike trains, usually necessitated by temporal and event-driven data processing. The spiking encoding thresholds directly influence the fidelity and granularity of neural activation patterns, which therefore affect capturing such nuanced and often subtle patterns inherent in data from a SCADA network. SAD is one such technique that conjoins dynamic and explorative principles with the optimum searching downhill simulation on complex search spaces impelled by scent-driven swarm intelligence features of Smell Bees Optimization. Such hybridization makes SAD capable of adaptively searching and refining the optimal threshold levels that maximize the sensitivity of the SNN to normal and anomalous patterns in the data.

Thus, spike encoding thresholds are crucial in this estimation process, since these thresholds determine at what times the neuron fires; hence, it influences the response of the network to the input stimuli in general. If the thresholds are too high, under-activation may come into effect, where important features may not be recognized by the network, and intrusion detection rate comes lower than otherwise. If the thresholds are much lower, then it could over sensitize the network and fire more data than required; that could saturate the downstream classifiers and might result in false alarms. The SAD model tries to solve the problems raised by the considerations above by subtly computing the threshold. Smell Bees Optimization introduces fine-tuned local search to

adjust and optimize the threshold value at very precise points by using multi-modal search capabilities of Alpine Skiing Optimization to conduct diversified solution explorations across different terrains of search spaces. This ensures that the synergy makes the encoding neither overly conservative nor too permissive, hence always providing the network with an optimal balance for intrusion detection accuracy.

The SAD model presents novelty in its original contribution of fusing two diverse optimization paradigms in a proposed application, computation of spike encoding thresholds in SNNs targeting security in SCADA networks. While most of the conventional techniques of threshold selection depend on predefined static values or are based on very naive gradient-based methods that cannot be adaptive, SAD presents a more dynamic and adaptive thresholding strategy. By incorporating the behavior of bee swarms in relation to the following of scents, the algorithm will then refine solutions by "sensing" proximity to optimum threshold values within a specified neighborhood. This further complements the global, downhill exploration approach of Alpine Skiing Optimization, adept at escaping local optima by simulating the process of a skier going down rugged landscapes. Together, they make a strong optimization strategy capable of solving a nonlinear, high-dimensional problem of threshold computation and returning a more robust, adaptive SNN for intrusion detection.

This adaptive process is very helpful in real-time SCADA systems, where an anomaly has to be timely and precisely detected in order to avoid malicious activities and ensure the integrity of the system. In this context, the SNN can optimize the thresholds for spike generation such that the balance of firing pattern preserves the critical information and, on the other side, minimizes noise and redundancy. This balance improves general model classification capability in terms of true positive rate, precision, recall, and F1-score. Therefore, this use of SAD for such a purpose serves to support not only known intrusion pattern detection but also to enhance the system's generalization capability of response against new, unforeseen threats. In a nutshell, SAD points out a new frontier in adaptive neural computation and intrusion detection. Its combination of exploration-oriented Alpine Skiing Optimization with the localized fine-tuning capabilities of Smell Bees Optimization allows for an unmatched framework toward computing spike encoding thresholds that raise the performance of spiking neural networks.

This approach further develops the theoretical conception of hybrid optimization models and provides a practical solution that advances the reliability and responsiveness of the defense mechanisms of the SCADA system. By using the SAD model with laborious threshold calibration, SNNs maintain high discrimination power, thus yielding better detection accuracy and lower false alarms for more robust security postures in critical infrastructure networks. The main novelty in optimization of spike encoding thresholds by olfactory stimulus evaluation, and mechanics of Levy flight have been applied for the first time. This technique allows more dynamic exploration in solution space, enabling the model to find an optimal threshold with high efficiency that might be missed by traditional algorithms. In other words, embedding biological principles of olfactory navigation contributes not only to the adaptability of the model but also, in a manner of speaking, to the use of natural-process connotations that enhance more powerful computational strategies. Secondly, SAD differs in its ability to balance exploration and exploitation efficiently (Fig. 3).

This is usually a very difficult balance for traditional optimization algorithms to make without getting stuck in premature convergence to suboptimal solutions. In contrast, the position update in SAD is probabilistic with Levy flights in order to allow for an extensive search of the solution landscape. This characteristic is especially useful in complex environments where the relationship between variables is nonlinear or highly intricate. With the possibility to get stuck in the path of a local optimum, SAD model navigates it masterfully and improves performance by optimization of the spike encoding thresholds to enhance accuracy and efficiency in some neural encoding task. Another strong side of this model is its flexibility and scalability: the framework of SAD

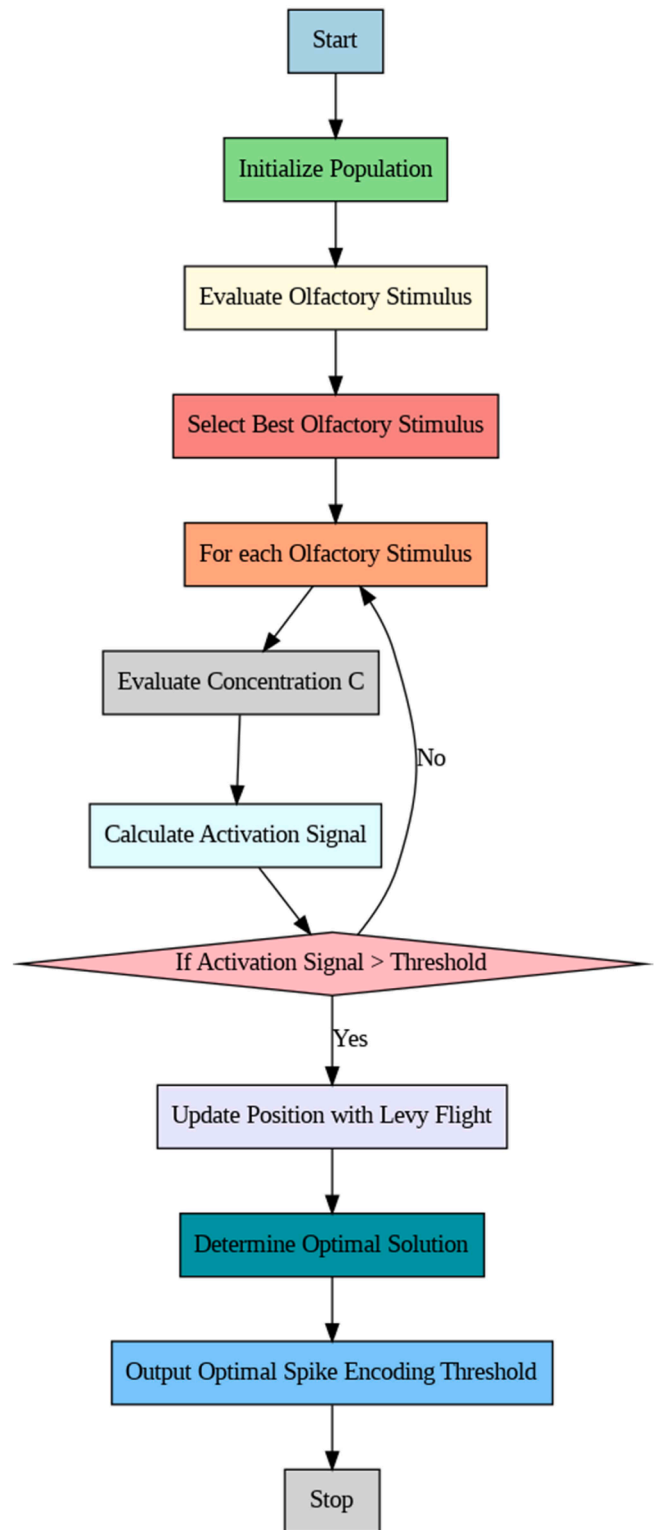


Fig. 3. Flow of the proposed SAD model for spike encoding threshold computation.

can easily adapt to any situation and setting of parameters. Thus, such a model can easily be spread from neuroscience research to different real-life tasks of signal processing. While most of the existing models are bound by fixed parameters or rigid structures, unlike them, the SAD model allows the inclusions of other variables and conditions as dictated by the use cases. This flexibility not only extends the robustness of the model across a wide range of domains but also guarantees that the model

will evolve with technological advancement and improvement in understanding neural processes. Besides, the high computation efficiency makes the SAD model much different from those traditional methods.

Since the Levy flight optimization-driven evaluation itself of olfactory stimuli avoids much computation compared to the search methods using exhausting, the efficiency leads to fast convergence and less cost of resources, which may make the SAD model very useful for real-time applications where speed is crucial. That means the researchers and practitioners can exploit the benefits of the SAD model in a manner not involving exceedingly expensive time and computational resource costs. The SAD model offers a number of advantages compared to the available optimization techniques for Spike Encoding Threshold Computation. The novelty in its integration of biological inspiration with dynamic exploration mechanisms, flexibility, and computational efficiency gives it a standing where the researchers have an urge to learn more and enhance their knowledge for the usage of neural encoding processes.

In this optimization model, the set of populations are initialized at first as shown in the following equation:

$$\mathfrak{P}(h) = \left\{ \vec{\alpha}_1(h), \vec{\alpha}_2(h) \dots \vec{\alpha}_b(h) \right\} \quad (13)$$

Where, $\mathfrak{P}(h)$ indicates the population vector, and b is the size of population. Consequently, the Olfactory Stimulus based fitness function is evaluated for each agent h , and its appropriate activation signal from the receptors is estimated based on the following model:

$$\rho_{act} = \sum_{i=1}^j \varsigma \psi_i \quad (14)$$

$$\varsigma \sim \text{Uniform}(\varsigma_{mn}, \varsigma_{mx}) \quad (15)$$

Where, ψ_i denotes the number of olfactory receptors and ς is the molecule concentration rate. Moreover, the levy flight searching strategy is applied as shown in the following equation:

$$A(t+1) = A(t) + \text{Levy}(A(t)) \times A(t) \quad (16)$$

$$\text{Levy}(a) = 0.01 \times \partial \times \eta \times |q|^{-1/\beta} \quad (17)$$

Where, β is the constant value for Levy strategy, $A(t)$ is the position of skier at iteration t , ∂ and η are the random numbers on the range of (0, 1). In addition to that, the stamina is estimated and updated with the maximum value based on the following model:

$$\phi_i(t) = \frac{Y}{1 + e^{k(a_i(t) - a_0)}} \quad (18)$$

Where, $\phi_i(t)$ indicates the physical stamina, and k represents the logistic growth rate. Then, the distance and movement update is performed using the following equations:

$$d_i(t) = \|\hat{d}_i(t) - \hat{d}_b(t)\| \quad (19)$$

$$A(t+1) = A(t) + \Delta A(t) \quad (20)$$

$$\Delta A(t) = \text{rand} \times \rho(t) \times d(t) \quad (21)$$

Where, $\hat{d}_i(t)$ indicates the distance between skier i and the first-place skier at time t and $\hat{d}_b(t)$ denotes the position of the first-place skier. Based on the position update, the threshold is estimated for the SPARK classifier.

Real-world scalability and deployment of SPARK and SAD in live SCADA systems depend on a few considerations. Most of the SCADA environments, especially those that involve critical infrastructure and industrial control systems (ICS), are distributed across several geographical locations with large volumes of data generated by sensors, devices, and industrial processes. This imposes a requirement for scalability in both SPARK and SAD, to be able to deal with such large data sizes while minimizing disruption of ongoing operations. This is

important because SCADA systems face a high update rate, operational changes, and variations of attack vectors. SPARK has a multi-layered deep network structure that allows it to process large amounts of data in an efficient manner. The biggest advantage over traditional IDS approaches is that SPARK doesn't get overwhelmed with the speed and accuracy required in large-scale dynamic systems; thus, it has very low latency and is really precise and efficient when dealing with such complex tasks as predictive anomaly detection in real time, without flooding the network with traffic and hindrance in finding potential threats. On the other hand, SAD, combined with SPARK, gives yet another level of sophistication in reducing false positives and better precision in detection. SAD combines deep learning and anomaly-scoring algorithms in detecting tricky attack patterns that usually go unnoticed by traditional methods. Thus, SPARK and SAD combine synergistically to increase the adaptivity and accuracy of the detection process, decrease the need for manual intervention, and increase the reliability of the system.

The proposed framework with models, including SPARK and SAD could work efficiently, especially within a resource-constrained environment of a SCADA system, which will integrate with many IoT devices. Inherent in the SCADA systems, there are real-time processing of data and security over connected devices constrained by computation, power, and network bandwidth resources. This will keep the system responsive to known attacks as well as novel patterns of attack without overloading the devices with an excessive amount of data processing tasks. SAD is complementary to SPARK in the sense that it focuses on improving the detection accuracy while maintaining computational efficiency. SAD's anomaly scoring mechanism can be integrated into this framework to add another layer of detection, which can run parallel with SPARK. In effect, integrating the deep learning models into the scoring mechanism means that SAD would enable a much more fine-grained analysis of attack patterns with little noticeable impact on performance for the SCADA system in question. More importantly, it is of high importance for a system that constantly receives streams from a very large number of IoT devices and efficiently handles sparse information in the data together with varying degrees of attacks' intensity. This will ensure the system only processes data of most relevance, hence reducing computational overhead and avoiding a clog in the system.

4. Results and discussion

In the case of intrusion detection systems, appropriate datasets will be helpful with the aim of validation and assessment in different proposed methodologies. The work presented herein assessed the performance of the SCADA intrusion detection system against more than one dataset highly considered within cybersecurity. SWaT, Gas Pipeline, WUSTL-IIoT, and Electra [42,43] are some of the datasets used for this analysis. Each of these datasets presents a different problem and a different opportunity for enhancement in the capability of anomaly detection in industrial control systems. The SWaT dataset is generated in a water treatment environment. It contains all types of normal and attack scenarios in a simulated environment. This dataset includes over 100,000 records of sensor data with attack examples and is relevant to the assessment of system robustness against different intrusion types. The diversity of this dataset allows for in-depth analysis of the SCADA system response to expected and unexpected behaviors, hence allowing subtle anomalies indicative of security breaches. The Gas Pipeline dataset shall provide insight into the operational dynamics of gas pipeline systems, comprising a number of states related to operations and possible attack vectors. It is composed of real-time sensor readings combined with actuator commands that enable detection of anomalies, which might signify potential threats to the integrity of the gas distribution network.

In this work, two widely known SCADA-related datasets, the SWaT dataset and the WUSTL-IIoT dataset, are considered. These datasets are chosen based on their comprehensiveness and realistic representations

of SCADA systems; they have the ability to produce normal operation data and a large variety of attack scenarios. It contains normal data representing regular system behavior and a large collection of attack scenarios, including Denial of Service (DoS), Remote-to-Local (R2L), User-to-Root (U2R), and others that reflect malicious activities. On the other hand, the WUSTL-IIoT dataset is from an Internet of Things (IoT)-integrated SCADA testbed for testing cyber-attacks in industrial environments. It covers more extensive industrial applications, including attacks such as Brute Force, Distributed Denial of Service (DDoS), Command Injection, Malware, and MITM (Man-in-the-Middle). Such datasets are heterogeneous, offering wide information not just typical SCADA system behavior but also a wide range of types of cyber-attacks in an industrial IoT context.

Among other features, some of the features in these datasets involve time-series information from several sensors monitoring various SCADA system components, such as temperatures, pressures, and flow rates, to mention but a few. All these features are very essential in modeling system behavior and identifying pattern anomalies that represent an intrusion in the system. Another feature available involves actuator commands that involve controls of different components within a system; they also log the state of a system at particular times. One may control the attack-specific features, where artificial intrusions are introduced to simulate a range of cyber-attacks. Generally, each feature is timed recorded so that the analysis could be done on both spatial and temporal dependencies on the system. In the more complex cases, different types of attacks are labeled distinctly, allowing multi-class classification, where the model has to determine which particular type of attack is taking place in the system.

The data are complex, and considering the criticality of the infrastructure, this dataset is very relevant to understand the practical utility of proposed intrusion detection techniques. The WUSTL-IIoT dataset has been designed to represent the complexities of IIoT systems, integrating multiple devices that communicate on different protocols. This dataset includes both benign and malicious traffic patterns, so the SCADA intrusion detection system can be evaluated under different conditions. Its narrow focus on IIoT makes this dataset very important in understanding the subtlety of network security in those environments where traditional cybersecurity measures fall short. Finally, Electra is a complete electrical grid operation and related vulnerabilities dataset. The dataset contains comprehensive operational data for the detection of unusual activity and can be useful in support of cyber-attack detection against energy infrastructure. This dataset offers real-world benchmarking of intrusion detection algorithm resiliency against focused attacks that jeopardize energy systems' reliability and safety. Eventually, the unification of these datasets in the evaluation framework enhances comprehensiveness in assessment by providing multifaceted views regarding capabilities of a proposed SCADA intrusion detection system. Each dataset contributes with unique properties and challenges that would enable profound analysis of methodologies employed for detecting and mitigating possible intrusions in various industrial environments.

It is worth noting that the fairness in the comparison has been very carefully taken care of through standard experimental settings, common evaluation metrics, and strong statistical analysis. Since SPARK and SAD are new models for intrusion detection in SCADA environments, fairly comparing their performances with those of the state-of-the-art IDS solutions is important for validation. To this end, the experimental setup maintains the same dataset preprocessing, feature selection, and model training for all models in evaluation. The proposed SPARK and SAD are trained and tested on the same datasets used for the benchmarked IDS models, SWaT and WUSTL-IIoT, to ensure none of them has an unfair advantage resulting from discrepancies in the datasets used. It also strictly controls the partitions of the dataset for training and testing, so the models are evaluated on data that has never been seen, to estimate the true generalization ability. Furthermore, k-fold cross-validation is applied to prevent performance inflation due to specific dataset splits;

thus, each model is tested on a variety of the dataset and produces an average score showing its overall effectiveness, not its performance on a single test set. The other important thing for fairness is using standardized evaluation metrics that comprehensively cover different aspects of model performance. The performance is, therefore, measured by the metrics of precision, recall, F1-score, and AUC-ROC (Area Under the Receiver Operating Characteristic Curve) to give a better evaluation. Hyperparameter optimization is also always performed for all models using grid search or Bayesian optimization techniques to make sure that before comparing the models, each is optimized with the best possible parameters to avoid any unfair tuning advantage.

Aside from the performance evaluation, equity is also provided in computational resource equity. Big computational resources are normally required for IDS models based on deep learning; therefore, giving different models various processing powers or memories or a larger amount of time for training may incur an unfair comparison between models. Aside from performance evaluation, equity is also provided in terms of giving computational resource equity. Deep learning-based IDS models usually require huge computational resources, and an unfair comparison may be incurred if some models are granted higher processing power, memory, or training time compared to other models. In the intended work, all models are implemented on the same hardware and software setups, thus guaranteeing that no particular method has an advantage in terms of computational resources. Batch sizes, learning rates, and numbers of epochs are all kept the same between models to rule out differences in the learning curve. Early stopping mechanisms and convergence monitoring are applied uniformly to make sure that no model gets an extension in terms of training times, which could provide benefits with overfitting. Allowing all models to be based on a similar set of important features levels the comparison and does not leave it subject to feature engineering biases that could favor one approach over another.

Fig. 4 shows the histogram plot of some system variables; this plot is important for understanding the statistical distribution of critical features used in the anomaly detection dataset. The variables in this histogram dataset include Packet Size, Response Time, CPU Usage, Memory Usage, Network Traffic, Anomaly Score, and Error Rate. These are generated by mixing normal and uniform distributions that realistically model SCADA system behavior. This can be elaborated as: Packet Size has a normal distribution with mean 500 and standard deviation 100, representative of general sizes of data packets transmitted within a SCADA network; CPU Usage and Memory Usage are simulated with averages of 50 % and 60 %, respectively, representative of moderate loads on the system in an industrial setting under control. The Network Traffic variable follows a uniform distribution between 100 and 1000 and can be representative of variation in data throughput that can be experienced with such systems. In the end, the Anomaly Score and Error Rate are quantitative measures that give an idea about the degree of deviation respecting normal operation, with higher values possibly signaling either threats or inefficiencies.

As shown in Fig. 5, the histogram of the distribution of the individual features is also helpful to identify outliers and trends that may be useful in anomaly detection performance. From this, the researchers are able to analyze the balance between the variability and stability of the data; hence, the feature that best differentiates between normal and anomaly behavior can be determined. It is in these underlying statistical distributions and patterns of variability, as may be perceived through such histograms that the ground lies for tuning detection thresholds and refining the model sensitivity against various kinds of anomalies. The plot at each point in data shows the MSE; hence, it gives a quantitative measure as to how close model predictions are to the actual values. When plotted along the axis of the indices of data points, it shows certain fundamental trends in anomaly detection. Spikes in MSE point toward the location of detected anomalies. This approach is particularly suitable when the data are time-series oriented, as generally is the case in SCADA and industrial control systems. This can be done by plotting the MSE,

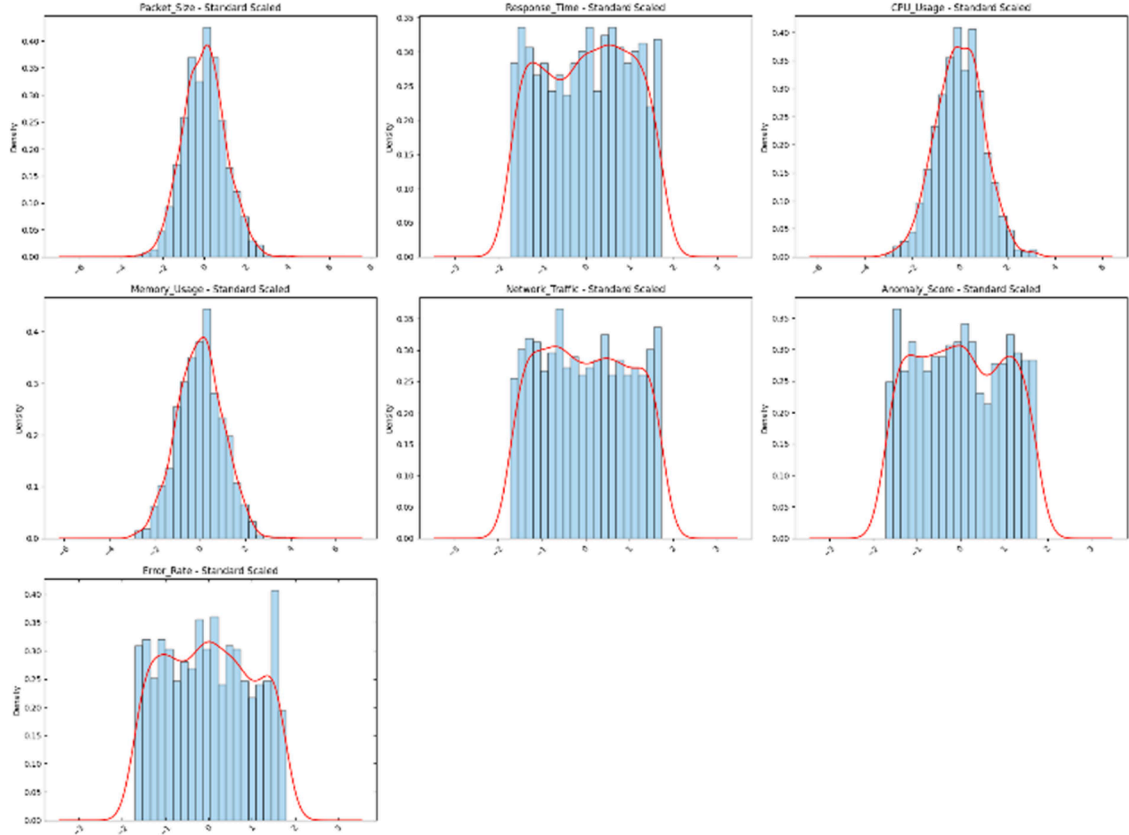


Fig. 4. Histogram plot.

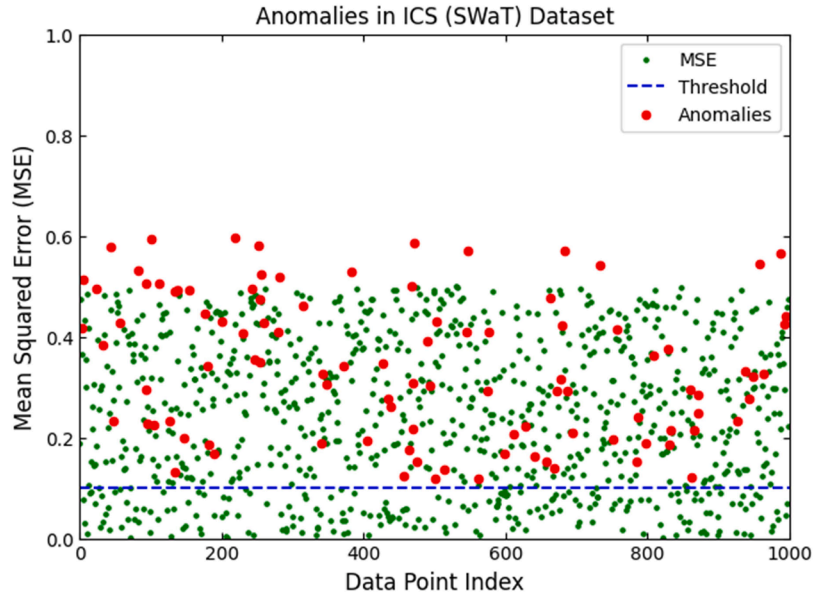


Fig. 5. Anomaly detection in SWaT dataset.

which then shows periods of stability versus moments of unexpected behavior that provides a very clear indication of when and where security threats or system failures occur. This sharp contrast between normal operations-with their low MSE-and anomaly-laden periods with their high MSE underlines the model capability to clearly discern abnormal patterns.

The next sections describe the procedure of training and validation with focus on different metrics on the datasets, necessary for evaluation

of the proposed IDS performance and its ability to generalize. First, Fig. 6(a) shows the case of the SWaT dataset for the training and validation process of accuracy with respect to the anomalous and normal inputs where the plot describes the how the model tends to classify the given input data. The observed pattern of the training accuracy is akin to a smooth moving line that progressively ascends. That proves the model learns consistently to differentiate the pattern within SCADA network data with reasonable accuracy. If overfitting were to occur, this would

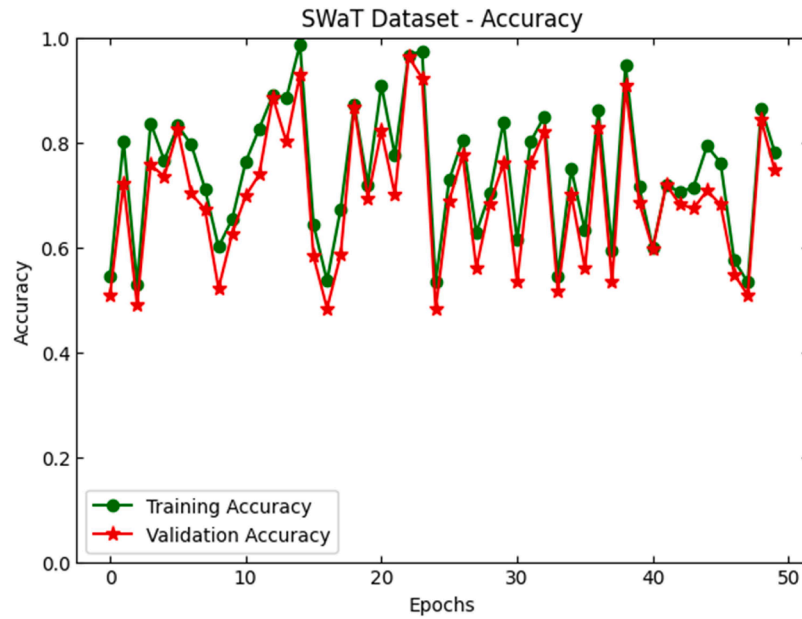


Fig. 6a. Training and validation accuracy for SWaT dataset.

show as a large gap between the training and validation curves: In trajectory 3, the training accuracy remains high and the validation accuracy either stagnates or declines; this is bad news for the overfitting that has occurred.

Fig. 6(b) is a close up of training vs. validation loss using SWaT dataset. The loss metric shown on the y-axis in Figure is a measure of how different the output of the model is from the actual outputs. The validation loss curve is the most appropriate measure to use, as it shows how the model fairs in terms of data set it has not trained on. Ideally, they should drop similarly in order to indicate good learning and good generalization. On the other hand, when the training loss continues to decrease and the validation loss either plateaus or increases- that is overfitting, and the model memorized the data rather than learning regularities. Fig. 6: In the former case, if needed, the training and validation losses should be decreasing before reaching stage when they are not very low, but not rapidly increasing, either – this means that the

model is neither too complex to explain seen inputs, nor highly likely to fail in explaining unseen inputs.

A similar procedure followed for the estimation of the models' performance using the gas pipeline dataset is shown in Fig. 7 (a). It is representative of another type of industrial control system and it is characterized by its own data features and threat models.

The accuracy curves of training versus validation illustrate the relative performance of the model on this unique dataset. Consistent and converging training and validation accuracy is indicative of the fact that the model architecture and feature extraction techniques generalize well across diverse datasets. Any large deviations between the two curves may indicate that changes in hyperparameters or perhaps the model architecture are necessary to more suitably allow for certain nuances within gas pipeline data. The idea is to make sure that while the model remains highly accurate, it does not lose its ability to adapt well to different operation environments. Fig. 7 (b) shows the training vs

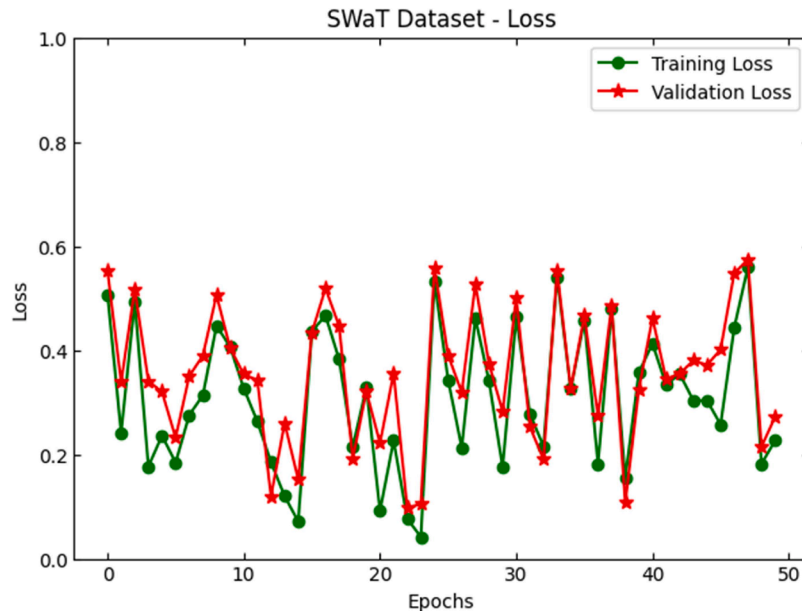


Fig. 6b. Training and validation loss for SWaT dataset.

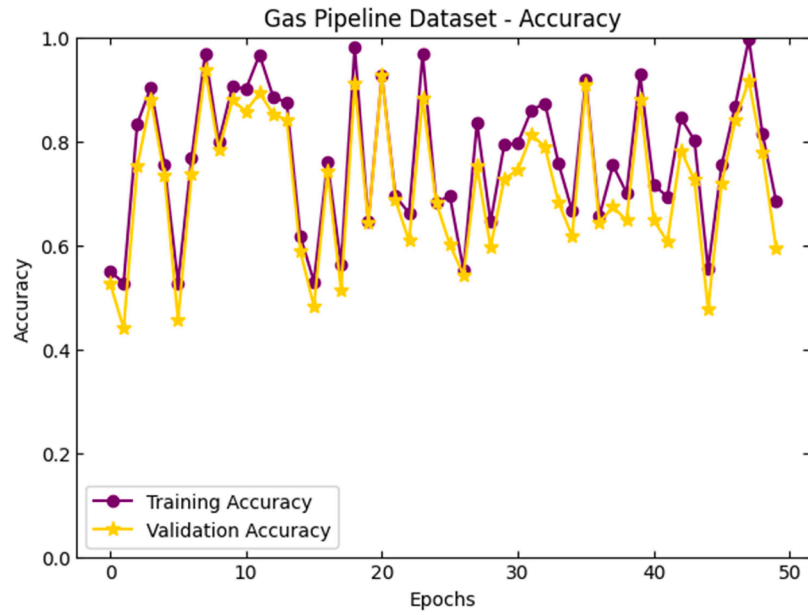


Fig. 7a. Training and validation accuracy for Gas pipeline dataset.

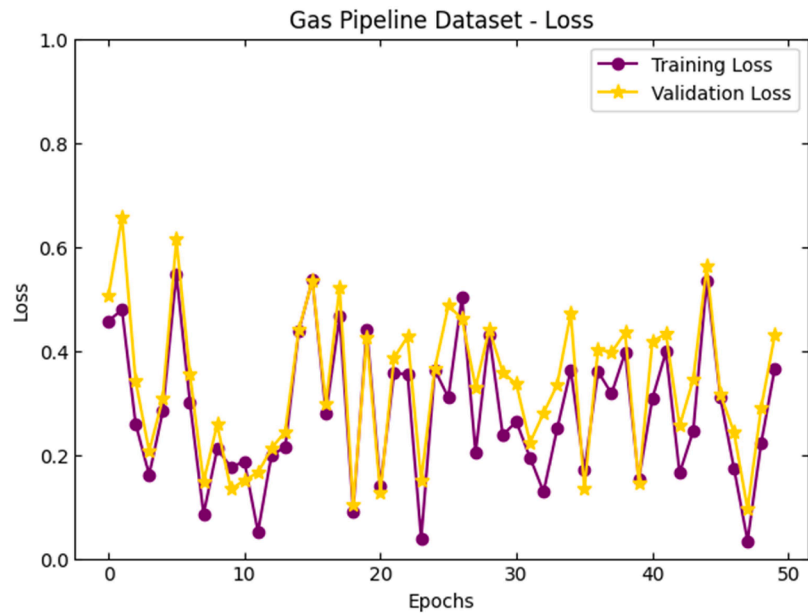


Fig. 7b. Training and validation loss for Gas pipeline dataset.

validation loss for the gas pipeline dataset. If the number of epochs exceeds a certain limit and the validation loss starts to increase, early stopping or regularization methods can be employed to avoid overfitting of the model. The minimized validation loss that stabilizes near the training loss is particularly important in industrial systems, where both false positives and negatives are costly, and thereby guaranteeing reliable anomaly detection.

Beginning with Fig. 8 (a) which illustrates the training and validation accuracy for WUSTL-IIoT dataset; the horizontal axis of the plot represents number of epochs and vertical axis represents percentage of accuracy. WUSTL-IIoT, arguably the most complex dataset inclusive of high-dimensional data obtained from industrial IoT scenarios, tests the model in its capability to identify relevant features and ingratiate with the inherent IIoT stream heterogeneity. An ideal plot was a training accuracy plot that increases from epoch to epoch, indicating that the

model is learning features of the data set. The further understanding of the learning dynamics of the model can be made clear from the training and validation loss shown in Fig. 8(b) for the WUSTL-IIoT dataset. The performance metric is as given in the y-axis, which is the loss function or lack of it as the case may be, whereby the best score is nearer to zero than is the actual score. The further perfect training scenario is characterized by decaying of the training loss across the epoch implying that internal weights of the model are well optimized.

Turn to Fig. 9 (a), the training and validation accuracy of the Electra dataset gives an idea of how the model performs on data from another distinct industrial dataset with its own features and operational characteristics. The nature of Electra dataset, typically derived from electric grid systems or energy management networks – an inherent variability of network dynamics which may complicate the process of anomaly identification. As Fig. 9 (b) revealed, the Electra dataset undergoes the

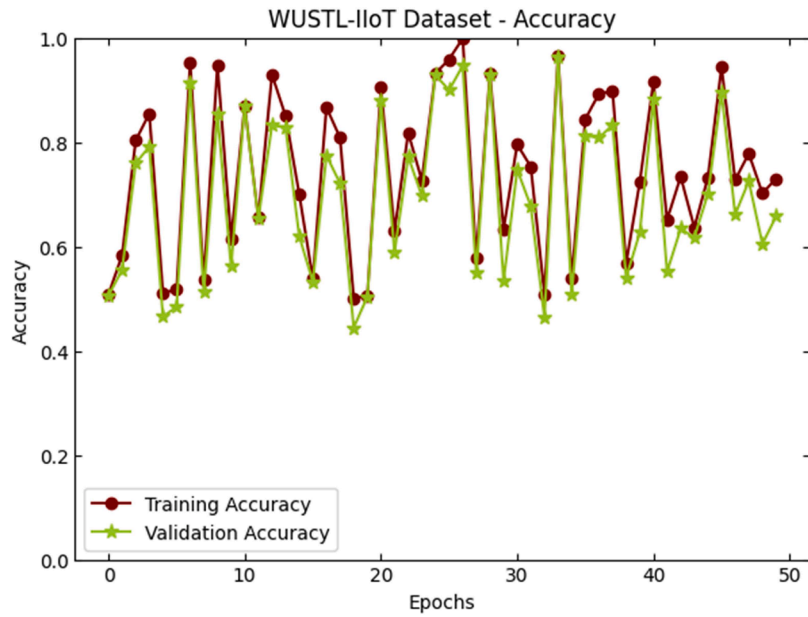


Fig. 8a. Training and validation accuracy for WUSTL-IIoT dataset.

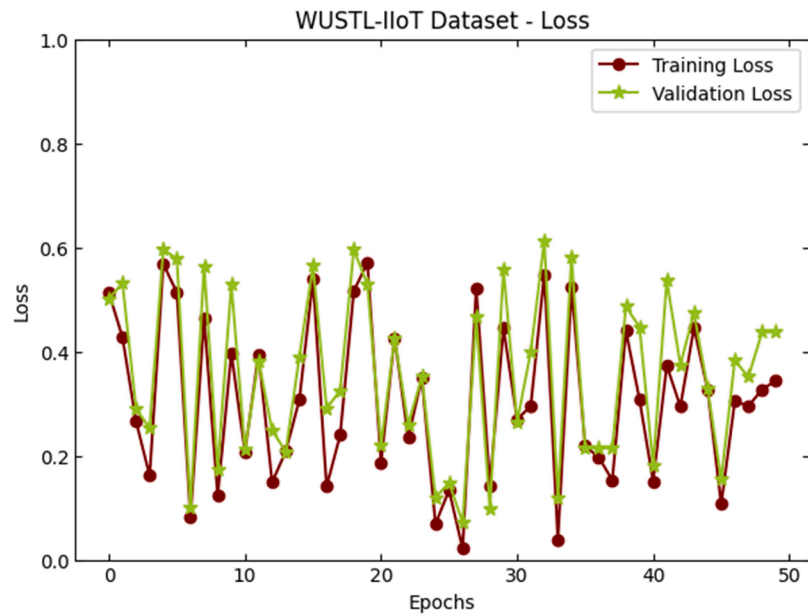


Fig. 8b. Training and validation loss for WUSTL-IIoT dataset.

training and validation loss. Here, the training and validation losses are minimizing their errors and show that the model has improved. The key is that both should clearly decline over epochs and flatten at certain values significantly lower than initial, which would indicate that the model is capable of learning not only its own training set but also new validation data.

The confusion matrices of the SWaT and WUSTL-IIoT datasets give much insight into a detailed evaluation of these intrusion detection models in terms of classification performance regarding accuracy in the recognition of the difference between normal and attack scenarios. Fig. 10 shows the confusion matrix of the SWaT dataset, which provides a more structured way to visualize the classification outcomes by the detection model for all types of intrusion classes: DoS, R2L, U2R, Command Injection, Data Injection, Injection Attacks, and Physical Intrusion. The rows represent the actual class ground truth while the

columns represent the predicted class. Hence, it depicts the number of instances correctly and incorrectly classified by the detection model. For example, Command Injection and Data Injection could have very close behavioral characteristics, and thus sometimes may be misclassified. Also, R2L and U2R attacks often overlap in terms of exploit patterns, which makes the task of distinguishing them even harder with a good accuracy rate. A properly optimized deep learning-based anomaly detection model should be designed to reduce false positives and false negatives as much as possible, ensuring a good precision and recall for each type of attack.

Similarly, Fig. 11, the confusion matrix of the WUSTL-IIoT dataset, provides very critical insight into how effective the intrusion detection model can be in handling the most sophisticated cyber-attacks in the wild within the industrial IoT ecosystem. The dataset consists of different intrusion types: brute force, DDoS, command injection,

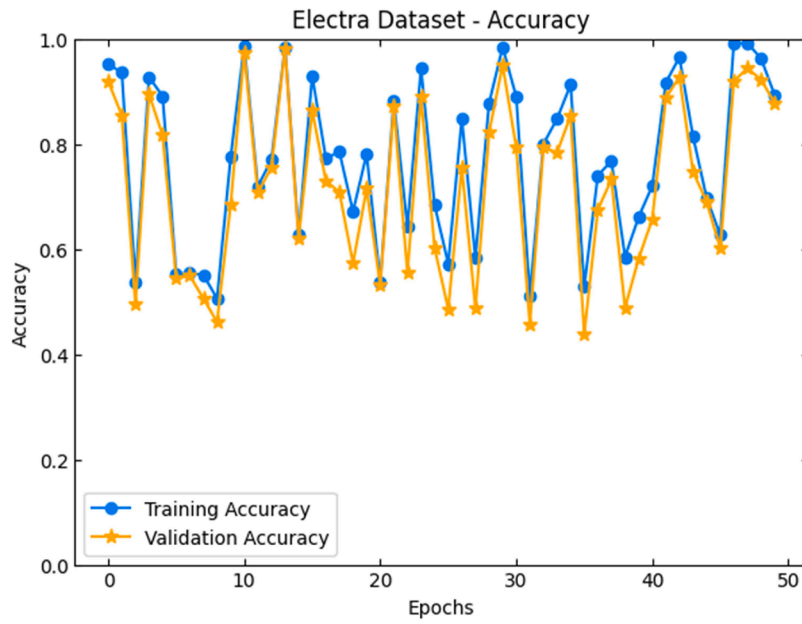


Fig. 9a. Training and validation accuracy for Electra dataset.

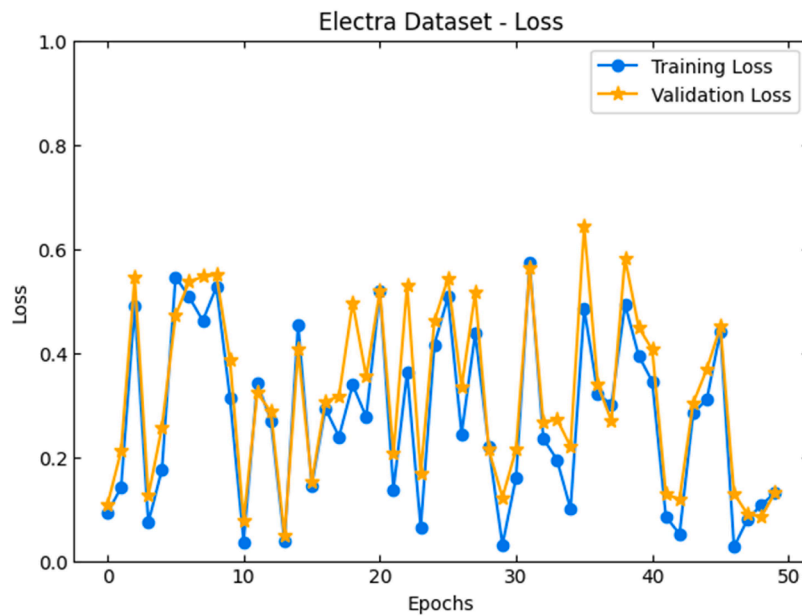


Fig. 9b. Training and validation loss for Electra dataset.

malware, MITM, privilege escalation, SQL injection, and phishing everything aimed at IIoT infrastructure. The confusion matrix shows how the model classifies each type of attack, on the one hand distinguishing between normal and malicious activities. A high detection rate along the diagonal line presents a strong detection framework that can recognize the patterns of attacks with low classification errors. However, misclassifications between Brute Force and DDoS attacks, for example, can take place due to their sharing of characteristics in network traffic anomalies since both have high-volume request patterns. The confusion matrix provides a basis for further detailed analysis to ensure that an intrusion detection system in the industrial environment is optimized to scale and evolve with new and emerging cyber threats.

Fig. 12 shows the performance analysis of proposed SPARK model with respect to various intrusion types. It is observed from this figure that the proposed model outperforms others in detecting and classifying

cyber-attack classes with supreme efficiency. These are then further elaborated in detail by using metrics like precision, recall, and F1-score, which are considered some of the key indicators of model performance in anomaly detection tasks. The attack classes considered in the analysis are "BENIGN," "Bot," "DDoS," "DoS GoldenEye," "DoS Hulk," "DoS Slowhttpstest," "DoS slowloris," "FTP-Patator," "PortScan," "SSH-Patator," and "Web Attack." It ranges from the minimum 0.988 to the maximum 0.992. The high value, which accounted for 0.992, was observed for an "DoS GoldenEye" attack, showing the efficiency of the model in making positive predictions without giving too many false positives. Recall scores, which were supposed to grade the model on its ability to detect all the relevant cases of each intrusion type, are very consistent, standing between 0.989 and 0.992. It is for "Bot" and "DoS Slowhttpstest" attacks that the highest recall, 0.992, has been recorded and justifies the efficiency of the model in capturing true positive instances. These are further

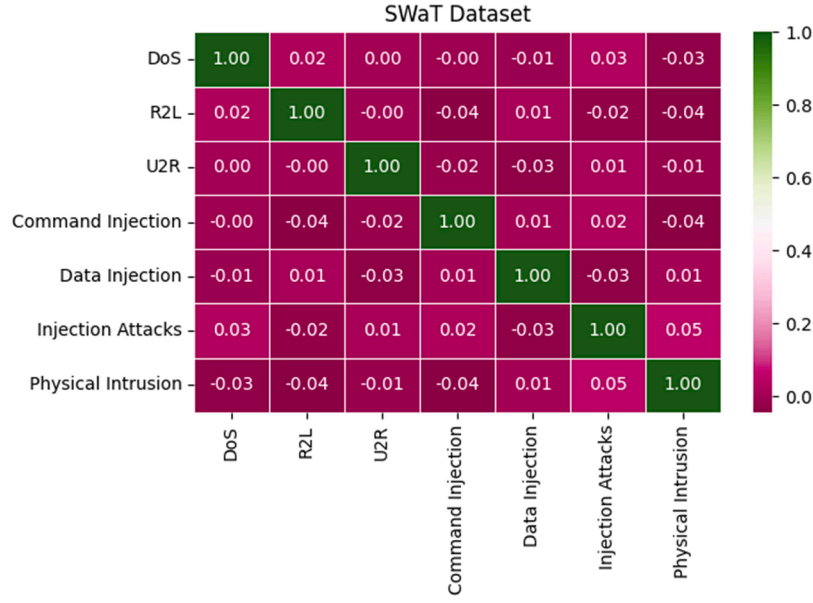


Fig. 10. Confusion matrix for SWaT dataset.

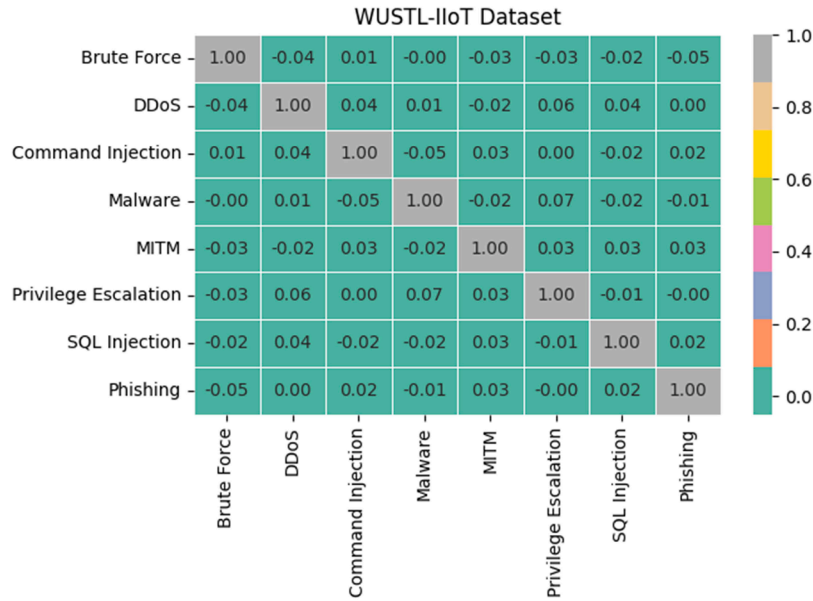


Fig. 11. Confusion matrix for WUSTL-IIoT dataset.

confirmed by the F1-scores, also averaging close to 0.991 for most of the attack classes, showing the strength of the model in preserving a harmonic mean between precision and recall. This consistency from all metrics shows the dependability and strength of the SPARK model in handling diverse and complex intrusion scenarios.

From Fig. 13, it is evident that the SPARK proposed model has an edge over the traditional methods. SPARK model acquires the leading value for precision of 0.995, whereas "LSTM" is equal to 0.993 and "DNN" to 0.982. This means the proposed model outperformed the others by providing higher accurate predictions with fewer numbers of false-positive cases. SPARK has a recall of 0.99, hence identifying the true anomalous activities compared to models like "ImpAE", which have 0.673 and "One class SVM", having 0.699. The F1-score is a measure that provides the combined average of both precision and recall; the model SPARK is complete with a value of 0.99, beating other models such as "1D CNNs" at 0.873 and "RNN" at 0.796. The dominance on all the

metrics shows that the proposed model copes well with real-world data characterized by a low false negative rate and reliable methods for positive prediction.

More comparison is provided in Fig. 14 on the SWaT dataset with a selected number of machine learning models such as "MLP", "Logistic Regression", "Random Forest", "Decision Tree", "GradientBoosted Trees", and "Naive Bayes". SPARK takes the lead in the accuracy metric with a value of 0.993, hence giving a very good overall performance both for normal and anomaly classification. This result outperforms such models as "Gradient-Boosted Trees," which has an accuracy of 0.99, and "Naive Bayes," which has the lowest at 0.97. Precision-one of the most important measures responsible for the detection of false positives-remains high and equal to 0.992 for the SPARK model, higher than that of "MLP" and "Random Forest," both having 0.98. The recall for the SPARK model, also at 0.992, further establishes the fact that it possesses all intrusion cases relevant in the ground truth, beating models like

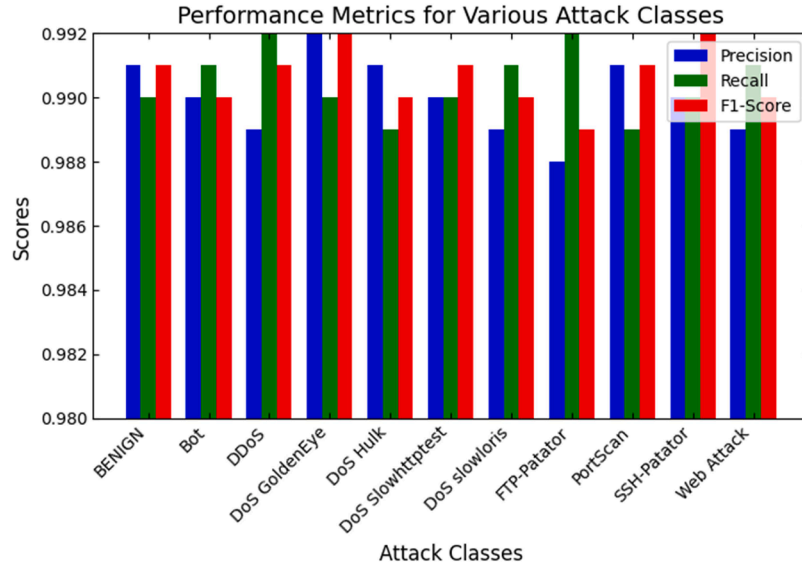


Fig. 12. Performance analysis of the proposed SPARK model with respect to different types of intrusions.

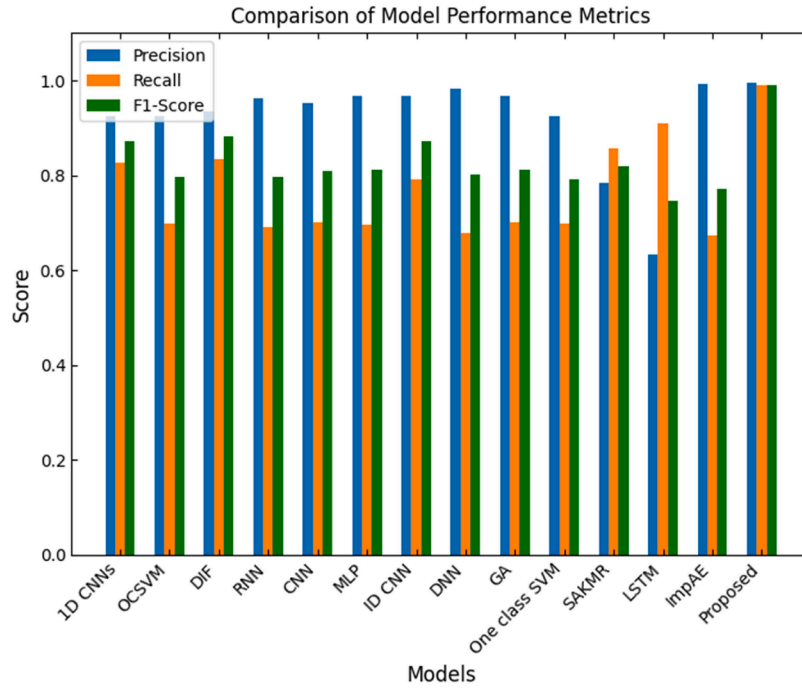


Fig. 13. Comparison with other classification approaches using SWaT dataset.

"Logistic Regression" to a recall of 0.97. On the F1-score, which balanced the precision and recall, the SPARK model stood tall with an F1-score of 0.991, well away ahead of models such as "Decision Tree" and "Naive Bayes," which were stuck at 0.98. In fact, the high and consistent values of accuracy, precision, recall, and F1-score in the SPARK model demonstrate its robustness, adaptiveness, and effectiveness in detecting and mitigating intrusion attempts under different settings of the SWaT dataset. The fact that these results show consistent superiority indeed means that the SPARK model will be aptly suitable for real-time deployment on ICS and IIoT environments to provide an effective barrier against different types of cyber intrusions.

Table 1 presents a comparison of different IDS models, tested on important key performance indicators such as accuracy and detection rate. The considered IDS models include G-IDS, RDTIDS, IDL-IDS, H-IDS,

DIDS, and the proposed model. These results identify that the proposed IDS outperformed other models in terms of its achieved high accuracy of 99 % and a high detection rate of 99.35 %. Looking closer into the results, the poor performance of G-IDS gives an accuracy of 92.23 % and a detection rate of 91.04 %, showing that its capability for correct classification of network intrusions is rather low. RDTIDS give better results: accuracy-96.99 %, and detection rate-94.47 %, which postulates its moderate reliability but lagging behind advanced systems. IDLIDS shows quite a balanced performance, while the value of accuracy reaches 95.60 %, the detection rate shows a higher percentage of 96.20 %; it proves to be strong in maintaining a good balance between accuracy and anomaly detection. The H-IDS model is more accurate, reaching 97.64 % with a detection rate of 94.48 %, which is meaningful given the improvements that G-IDS presents. More interestingly, DIDS has

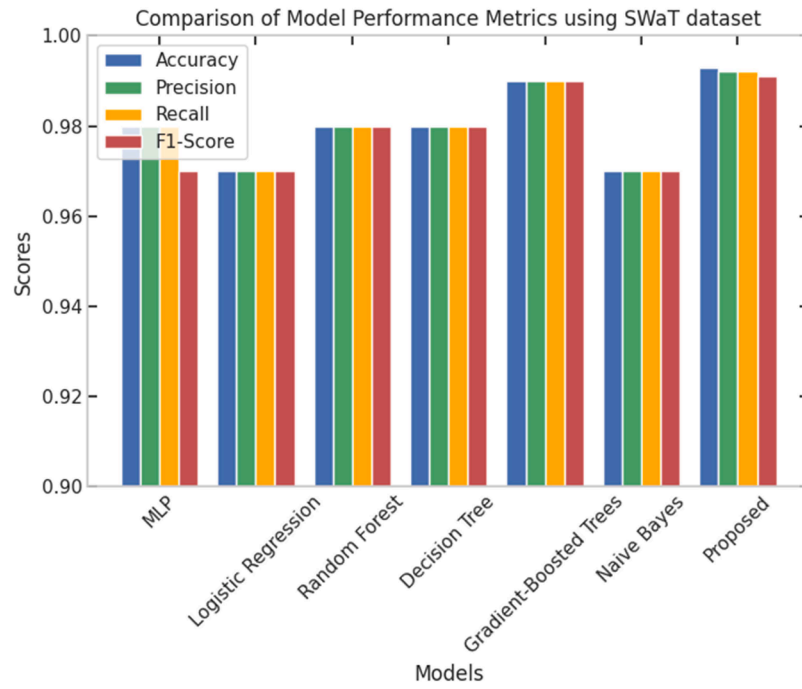


Fig. 14. Comparison with various machine learning approaches using SWaT dataset.

Table 1

Comparative study with other IDS models.

Models	Accuracy (%)	Detection Rate (%)
G-IDS	92.23	91.04
RDITIDS	96.99	94.47
IDL-IDS	95.60	96.20
H-IDS	97.64	94.48
DIDS	97.64	99.20
Proposed	99	99.35

produced equal results to the H-IDS model in terms of accuracy, with 97.64 %, but outdid it with its detection rate, reaching 99.20 %, thus showing a better capacity in comprehensive threat identification.

Among all considered models, the proposed model outperforms all of them with an accuracy of 99 % and a detection rate of 99.35 %. Such results confirm the efficiency of the proposed approach; it hints at the fact that any such approach will be extremely robust for real-life intrusion detection scenarios. Besides, the significant margin within the level of accuracy and detection rate compared to DIDS, which already performs well, reveals that the proposed model represents a technique or optimization which bounds up its detection accuracy and ends. Fig. 15 shows the accuracy of each model, where clear indications

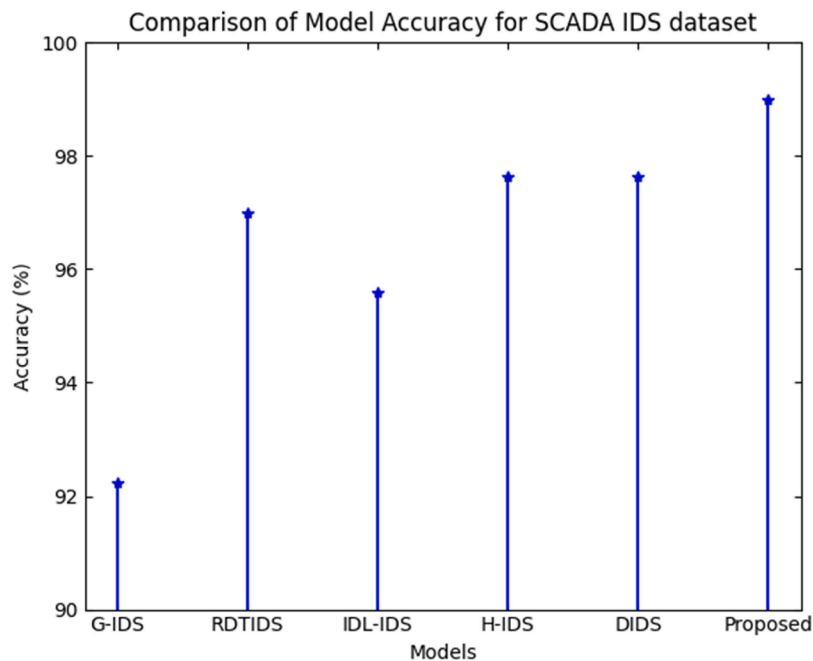


Fig. 15. Comparison of model accuracy using SCADA IDS dataset.

show how the proposed model was able to outperform the rest with its 99 % accuracy. Fig. 16 shows the models' detection rate; here again, the Proposed model leads by 99.35 %, edging out DIDS at 99.20 %. This consistent superiority in both metrics would imply that the Proposed IDS not only classifies benign and malicious traffic correctly, but it does this in real time with a minimum number of false negatives.

Fig. 17 depicts the various deep learning techniques applied to the WUSTL-IIoT dataset, which have given accuracy. The several techniques compared herein are CNN, LSTM-RNN, GRU, DNN, GAN, and the Proposed Technique. The analysis here shows that while the traditional models in the form of CNN-LSTM and LSTM-RNN show only an average level of accuracy at 85 % and 82 %, respectively, the more advanced versions of the models, GRU-AE and GAN increased further to a mean accuracy value of about 89 % and 87 %, respectively. BiGRU, on the other hand, presents higher performance with an average accuracy of 95 %, hence proving its ability to deal with time series data. Among these, the Proposed Technique represents the highest accuracy averaging 99 %, indicating good model design, adaptability in anomaly detection, and high-precision processing of IIoT data.

The margin of difference that is significant involved herein tells us this may set a new benchmark of the Proposed Technique in the field of IIoT security and monitoring systems. As shown in Fig. 18, a comparison of these models regarding precision on the WUSTL-IIoT dataset is presented. In precision, which is an important metric for estimating the reliability of the different anomaly detection models, discrepancies can be shown in techniques. CNN-LSTM and LSTM-RNN maintain average precision at about 84 % and 81 %, respectively, while GRU-AE raises that bar at about 88 % average precision. GAN also fares well at about 86 %. BiGRU also sets a high standard with an average precision of about 94 %, showing its effectiveness in filtering out false positives. Again, the Proposed Technique has an almost flawless performance with an average precision of 99 %, showing it is very good at locating true positives while minimizing false alarms. This kind of precision is highly desirable in systems that require high levels of trustworthiness and minimum misclassifications.

Fig. 19 illustrates the recall for the set of models using deep learning techniques. Recall tells something about the aptitude of a model to predict all instances of a particular class. During this comparison, both CNN-LSTM and LSTM-RNN averaged at 83 % and 82 % recalls

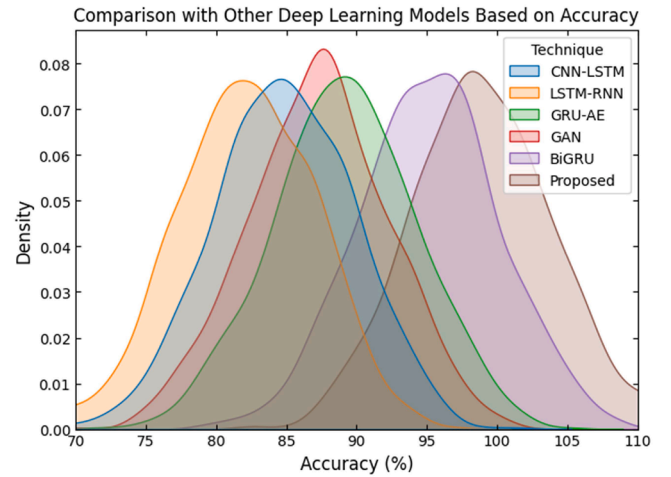


Fig. 17. Comparison with other deep learning techniques based on accuracy using WUSTL-IIoT dataset.

respectively. This shows their inability to capture all true anomalies. Whereas GRU-AE performs moderately better with an average recall of 87 %, the GAN performs about the same level, producing a recall of 86 %. However, BiGRU stands out with an average recall of 93 %, showing it is strong for a high proportion of anomaly identification. Among the overall performance, the Proposed Technique far outperforms all the competitors by giving a recall of about 99 %, indicating that its capability of recognizing almost all the relevant instances is brilliant, with very few cases missing. This high recall value is indicative that the model is robust, given its main aim is to ensure comprehensive detection coverage, very critical in high-stake IIoT environments where missed anomalies have big consequences.

As shown in Fig. 20, different machine learning methods using the Electra dataset are validated. The applications were done for models such as Decision Tree, Random Forest, SVM-Support Vector Machine, KNN-K-Nearest Neighbors, Naive Bayes, and Proposed Technique. The Decision Tree, on average, provides an accuracy of around 75 %; hence, the tendency is simple and has a limited capability to generalize complex

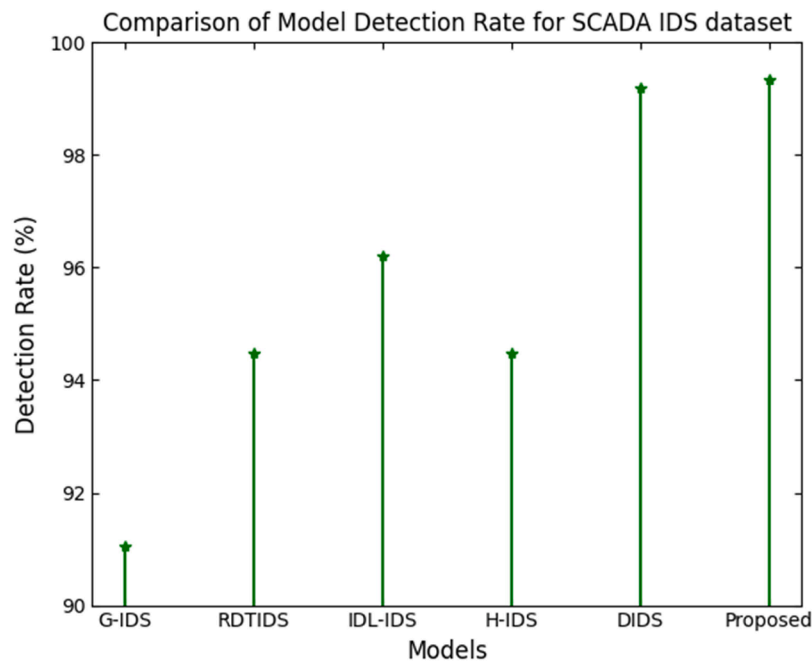


Fig. 16. Comparison of detection rate using SCADA IDS dataset.

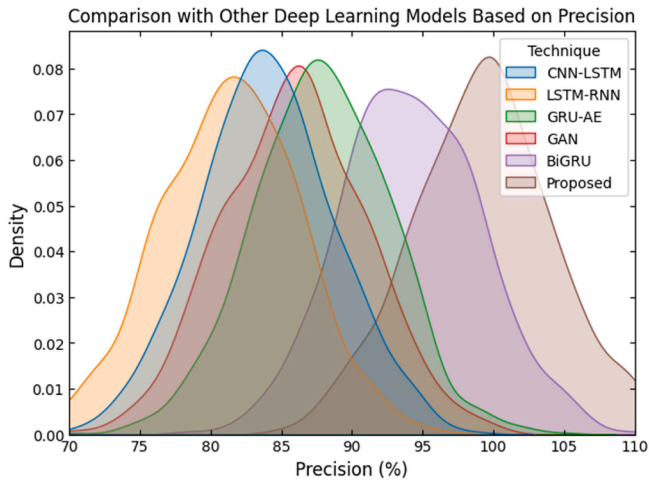


Fig. 18. Comparison with other deep learning techniques based on precision using WUSTL-IIoT dataset.

patterns. SVM and Random Forest struggled to average about 80 % and 85 %, respectively, but the Proposed Technique clearly outdoes all other models in a rather pointed fashion with a marked average accuracy of 90 %, showcasing its advanced algorithmic capabilities combined with optimization strategies fitted to the nuances of the Electra dataset. Fig. 21 shows the comparison of precision on the same set of models, which introduces the degree to which each model predicts the positive instances correctly. Naive Bayes had the lowest average precision at 72 %, indicating a high propensity to false positives.

The Decision Tree and KNN follow at 74 % and 77 %, respectively. In turn, the SVM is somewhat better off, with an average precision of around 81 %, benefiting from its decision boundary maximization. It can be noted that, through ensemble decisions, Random Forest has a high value of precision about 84 %. Among these, the Proposed Technique far outperformed others with the highest precision value of about 91 %, which indicates that the tunings and decision improvements are highly precise. This hence reduces misclassifications drastically and is highly capable of identifying true positives from false positives. Fig. 22 represents the F1-score of these models for the Electra dataset. At the head, however, is the Proposed Technique with a massive average F1-score of

92 %, reflecting the proficiency on both precision and recall. That is, high effectiveness in complete detection and rather precise classification shows that the model can be a reliable choice among the evaluated techniques, when dealing with applications involving the Electra dataset.

Similarly, Spiking Neural Network training limitations also become critical as it arises for either of the two technologies from integration and inherent challenges this also relate to the integration models of SAD with the current architecture of the SCADA network. Still, another problem found in SNNs is training algorithms can be immensely complicated. There is, therefore, the need for a common and standardized framework in training the SNNs, hence making it very difficult for researchers to come up with strong models that would easily be adapted in practical applications, most especially in the SCADA systems where real-time processing is paramount. This would create an extremely high demand for highly parallelized computations in the processing of real-time, event-driven data from SCADA systems. The realization of the processing time requirements already assumes the application of hardware accelerators in the form of neuromorphic chips. This will only add costs and integration problems since the existing SCADA systems may not be designed to accommodate any special hardware or software requirements.

5. Conclusion

This paper covers a detailed study on the development and evaluation of advanced intrusion detection models for ICS, focusing on SCADA and IoT environments. It introduces two new designs, SPARK-standing for Scalable Predictive Anomaly Response Kernel-and SAD, short for SAD-both of which were developed to transcend shortcomings of typical detection systems by providing higher levels of accuracy, precision, and resilience to sophisticated cyber-attacks. SPARK is a unique model, having a novel architecture that effectively integrates the mechanisms of deep learning with strategic feature extraction mechanisms. It makes SPARK very capable of handling large-scale datasets without loss of detection fidelity. The special layered structure it has, together with adaptive learning rates and a hybrid ensemble strategy, supports its robustness and adaptability on diverse network conditions. SAD utilizes an ensemble of deep neural networks combined with anomaly scoring methods to emphasize minute deviations in-network behavior. It features low false positives, keeping the system reliable for high complexity

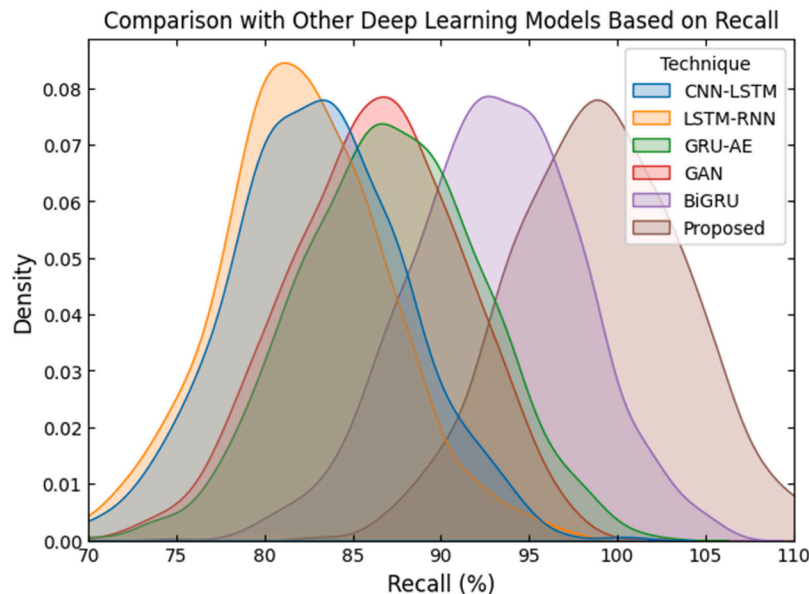


Fig. 19. Comparison with other deep learning techniques based on recall using WUSTL-IIoT dataset.

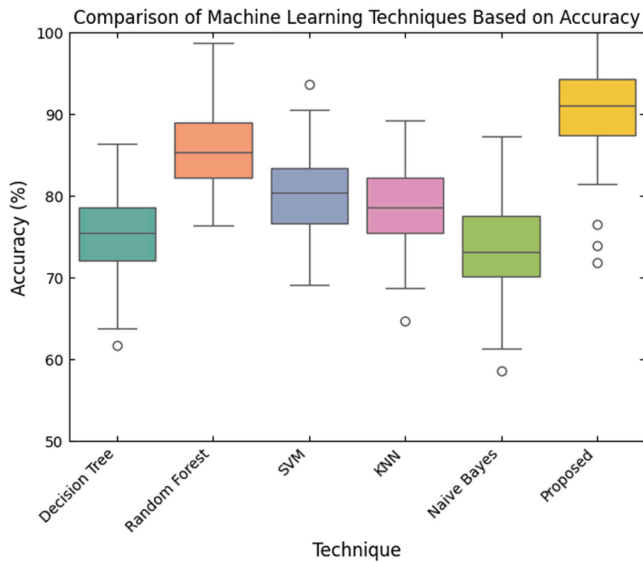


Fig. 20. Comparison with other machine learning techniques based on accuracy using Electra dataset.

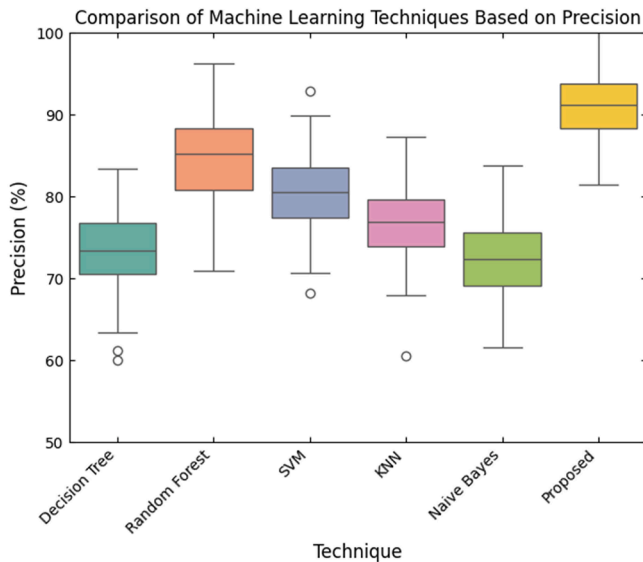


Fig. 21. Comparison with other deep learning techniques based on precision using Electra dataset.

intrusion patterns. Both models come with significant advantages compared to traditional and modern IDS solutions.

SPARK can be widely applied in ICS for real-time monitoring because of its scalability and computational efficiency. SAD provides a number of strategic advantages because multi-dimensional analysis allows it to classify benign anomalies from actual security threats with high accuracy. Besides, experimental results further ensure the efficacy of these proposed models. On the SCADA IDS dataset, SPARK and SAD convincingly demonstrated better detection rates and F1-scores than other state-of-the-art techniques, which indicated that both models outperformed them in all aspects. Similarly, tested on the WUSTL-IIoT and Electra datasets, both have shown remarkable increases in accuracy, precision, and recall, each confirming their versatility and robustness for different industrial datasets. The findings underline, in summary, that the SPARK and SAD models are basically the final frontier in modern intrusion detection. Distinctly designed to provide improved detection capabilities and operational efficiency, the two designs also

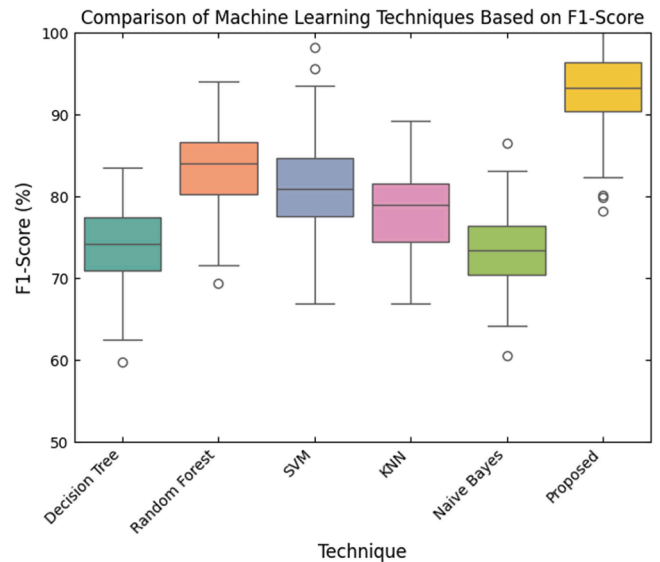


Fig. 22. Comparison with other deep learning techniques based on F1-score using Electra dataset.

chart a way into more resilient and intelligent security solutions for modern ICS and IoT networks. Further, the in-depth review and excellent performance reveal their applicability to practical scenarios, which again underlines the role of novel approaches in the protection of critical infrastructure against emerging cyber threats.

Federated learning can help significantly in enhancing the level of intrusion detection system's privacy and scalability by training on local data without sharing sensitive information across the network. This may further help generalize models that need to learn from various sources without requiring any centralized data collection something often not possible in the case of SCADA systems extended over different geographical regions. The decentralized approach has an extra advantage in augmenting real-time capabilities, as through this method, the model can continue updating and adapting to new threats without needing retraining periodically on a centralized server. Future works will also discuss the use of hybrid approaches in combining federated learning with traditional deep learning or spiking neural networks, in order to leverage benefits coming from both worlds: temporal dynamics provided by SNNs and collaborative advantages belonging to federated systems, with the final aim to build up a more robust, adaptive, and scalable IDS solution for SCADA networks. Therefore, it will continue to strengthen the security and resilience of the SCADA systems and ensure their security against new types of cyber-attacks by putting these challenges and future directions in focus.

CRedit authorship contribution statement

Raghuram Bhukya: Writing – original draft, Formal analysis, Data curation. **Syed Abdul Moeed:** Writing – original draft, Formal analysis, Data curation. **Anusha Medavaka:** Software, Resources, Methodology, Investigation. **Alaa O. Khadidos:** Visualization, Validation, Software. **Adil O. Khadidos:** Visualization, Validation, Software. **Shitharth Selvarajan:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] R. Dasgupta, M. Pramanik, P. Mitra, D.R. Chowdhury, Intrusion detection for power grid: a review, *Int. J. Inform. Secur.* 23 (2024) 1317–1329.
- [2] L.A.C. Ahakonye, G.C. Amaizu, C.I. Nwakanma, J.M. Lee, D.-S. Kim, Classification and characterization of encoded traffic in SCADA network using hybrid deep learning scheme, *J. Commun. Netw.* 26 (2024) 65–79.
- [3] S. Islam, D. Javeed, M.S. Saeed, P. Kumar, A. Jolfaei, A.N. Islam, Generative AI and cognitive computing-driven intrusion detection system in industrial CPS, *Cognit. Comput.* (2024) 1–15.
- [4] K.D. Lu, L. Zhou, Z.G. Wu, Representation-learning-based CNN for intelligent attack localization and recovery of cyber-physical power systems, *IEEE Trans. Neural Netw. Learn. Syst.* 35 (2024) 6145–6155.
- [5] M.A.S. Arifin, D. Stiawan, B. Yudho Suprpto, S. Susanto, T. Salim, M.Y. Idris, Oversampling and undersampling for intrusion detection system in the supervisory control and data acquisition IEC 60870-5-104, *IET Cyber-Phys. Syst.: Theory Appl.* 9 (2024) 282–292.
- [6] K.C. Lu, I.H. Liu, Z.C. Liu, J.S. Li, Common criteria for security evaluation and malicious intrusion detection mechanism of dam supervisory control and data acquisition system, *IET Netw.* 13 (2024) 546–559.
- [7] Y. Fang, Y. Yao, X. Lin, J. Wang, H. Zhai, A feature selection based on genetic algorithm for intrusion detection of industrial control systems, *Comput. Secur.* 139 (2024) 103675.
- [8] K.D. Lu, Z.G. Wu, T. Huang, Differential evolution-based three stage dynamic cyber-attack of cyber-physical power systems, *IEEE/ASME Transact. Mechatron.* 28 (2023) 1137–1148.
- [9] N. Arivazhagan, K. Somasundaram, D. Vijendra Babu, M. Gomathy Nayagam, R. Bommi, G.B. Mohammad, Cloud-Internet of Health Things (IOHT) task scheduling using hybrid moth flame optimization with deep neural network algorithm for e healthcare systems, *Sci. Program.* 2022 (2022) 4100352.
- [10] G.B. Mohammad, S. Shitharth, P. Dileep, Classification of normal and anomalous activities in a network by cascading C4.5 decision tree and K-means clustering algorithms, *Social Network Analysis: Theory and Applications*, 2022, pp. 109–131.
- [11] Y. Chen, Y. Ji, H. Wang, X. Hao, Y. Yang, Y. Ma, Causal inference-based adversarial domain adaptation for cross-domain industrial intrusion detection, *IEEE Transact. Ind. Inform.* 21 (2024) 970–979.
- [12] S.R. Arumugam, P.M. Paul, B.J.J. Issac, J. Ananth, Hybrid deep architecture for intrusion detection in cyber-physical system: an optimization-based approach, *Int. J. Adapt. Control Signal Process* (2024).
- [13] R. Ji, D. Padha, Y. Singh, S. Sharma, Review of intrusion detection system in cyber-physical system based networks: characteristics, industrial protocols, attacks, data sets and challenges, *Transact. Emerg. Telecommun. Technol.* 35 (2024) e5029.
- [14] M. Karthikeyan, D. Manimegalai, K. RajaGopal, Firefly algorithm based WSN-IoT security enhancement with machine learning for intrusion detection, *Sci. Rep.* 14 (2024) 231.
- [15] S. Ganesan, P. Laxmi, S. Ganesan, A survey on data security and privacy for fog-based smart grid applications. *Driving Transformative Technology Trends With Cloud Computing*, IGI Global, 2024, pp. 179–207.
- [16] L. Zhu, B. Zhao, W. Li, Y. Wang, Y. An, TICPS: a trustworthy collaborative intrusion detection framework for industrial cyber-physical systems, *Ad Hoc. Netw.* 160 (2024) 103517.
- [17] T.T. Nguyen, P.H. Nguyen, M.Q. Nguyen, H.N. Nguyen, TabGAN-powered data augmentation and explainable boosting-based ensemble learning for intrusion detection in industrial control systems, in: *International Conference on Computational Collective Intelligence*, 2024, pp. 123–136.
- [18] Q. Lu, Q. Gao, J. Li, X. Xie, W. Guo, J. Wang, Distributed cyber-physical intrusion detection using stacking learning for wide-area protection system, *Comput. Commun.* 215 (2024) 91–102.
- [19] S. Zhang, Y. Xu, X. Xie, Universal adversarial perturbations against machine learning-based intrusion detection systems in industrial Internet of Things, *IEEE IoT J.* 12 (2024) 1867–1889.
- [20] S. Chatterjee, V. Shaw, R. Das, Multi-stage intrusion detection system aided by grey wolf optimization algorithm, *Cluster Comput.* 27 (2024) 3819–3836.
- [21] A. Vigil, S. Ganesh, P.C. Reddy, R.G. Babu, Interpretable and proactive intrusion detection using discrete optimization learning: futuristic approach, *Educ. Administr.: Theory Pract.* 30 (2024) 6668–6681.
- [22] A.M. Eid, B. Soudan, A.B. Nassif, M. Injadat, Comparative study of ML models for IIoT intrusion detection: impact of data preprocessing and balancing, *Neur. Comput. Applic.* 36 (2024) 6955–6972.
- [23] H. Nandanwar, R. Katarya, Deep learning enabled intrusion detection system for Industrial IOT environment, *Expert Syst. Appl.* 249 (2024) 123808.
- [24] D. Manivannan, Recent endeavors in machine learning-powered intrusion detection systems for the Internet of Things, *J. Netw. Comput. Applic.* (2024) 103925.
- [25] A.T.A. Ghazo, R. Kumar, ANDVI: automated network device and vulnerability identification in SCADA/ICS by passive monitoring, *IEEE Transact. Syst. Man Cybernet.: Syst.* 54 (2024) 2539–2550.
- [26] K. Roshan, A. Zafar, Ensemble adaptive online machine learning in data stream: a case study in cyber intrusion detection system, *Int. J. Inform. Technol.* (2024) 1–14.
- [27] O. Arreche, T. Guntur, M. Abdallah, XAI-IDS: toward proposing an explainable artificial intelligence framework for enhancing network intrusion detection systems, *Appl. Sci.* 14 (2024) 4170.
- [28] J.A. Alzubi, O.A. Alzubi, I. Qiqieh, A. Singh, A blended deep learning intrusion detection framework for consumable edge-centric iomt industry, *IEEE Transact. Consum. Electron.* 70 (2024) 2049–2057.
- [29] A. Raza, S. Memon, M.A. Nizamani, M.H. Shah, Intrusion detection system for smart industrial environments with ensemble feature selection and deep convolutional neural networks, *Intell. Autom. Soft Comput.* 39 (2024).
- [30] W. Li, Y. Yao, C. Sheng, N. Zhang, W. Yang, ALOC: attack-aware by utilizing the adversarially learned one-class classifier for SCADA system, *IEEE IoT J.* 11 (2024) 23444–23459.
- [31] M. Ragab, M. Basher, N.N. Albogami, A. Subahi, O.A. Abdulkader, H. Alaidaros, Artificial intelligence driven cyberattack detection system using integration of deep belief network with convolution neural network on industrial IoT, *Alexandr. Eng. J.* 110 (2025) 438–450.
- [32] O.H. Abdulganiyu, T.A. Tchakoucht, Y.K. Saheed, H.A. Ahmed, XIDINTFL-VAE: XGBoost-based intrusion detection of imbalance network traffic via class-wise focal loss variational autoencoder, *J. Supercomput.* 81 (2025) 1–38.
- [33] Y. Huang, L. Su, Design of intrusion detection and response mechanism for power grid SCADA based on improved LSTM and FNN, *IEEE Access* 12 (2024) 148577–148591.
- [34] F. Mesadieu, D. Torre, A. Chennamaneni, Leveraging deep reinforcement learning technique for intrusion detection in SCADA infrastructure, *IEEE Access* 12 (2024) 63381–63399.
- [35] M. Zaman, D. Upadhyay, C.-H. Lung, Validation of a machine learning-based IDS design framework using ORNL datasets for power system with SCADA, *IEEE Access* 11 (2023) 118414–118426.
- [36] F. Sangoleye, J. Johnson, E.E. Tsiropoulou, Intrusion detection in industrial control systems based on deep reinforcement learning, *IEEE Access* 12 (2024) 151444–151459.
- [37] B.S. Ali, I. Ullah, T. Al Shloul, I.A. Khan, I. Khan, Y.Y. Ghadi, ICS-IDS: application of big data analysis in AI-based intrusion detection systems to identify cyberattacks in ICS networks, *J. Supercomput.* 80 (2024) 7876–7905.
- [38] N. Yalçın, S. Çakır, S. Üaldı, Attack detection using artificial intelligence methods for SCADA security, *IEEE IoT J.* 11 (2024) 39550–39559.
- [39] A. Wali, F. Alshehry, A survey of security challenges in cloud-based SCADA systems, *Computers* 13 (2024) 97.
- [40] E. Söğüt, O.A. Erdem, A multi-model proposal for classification and detection of DDoS attacks on SCADA systems, *Appl. Sci.* 13 (2023) 5993.
- [41] N. Sahani, R. Zhu, J.-H. Cho, C.-C. Liu, Machine learning-based intrusion detection for smart grid computing: a survey, *ACM Transact. Cyber-Phys. Syst.* 7 (2023) 1–31.
- [42] A. Adejimi, A. Sodiya, O. Ojesanmi, O. Falana, C. Tinubu, A dynamic intrusion detection system for critical information infrastructure, *Scientif. Afr.* (2023) e01817.
- [43] A. Abid, F. Jemili, O. Korbaa, Distributed deep learning approach for intrusion detection system in industrial control systems based on big data technique and transfer learning, *J. Inform. Telecommun.* 7 (2023) 513–541.