Securing and Optimizing Wireless Sensor Military Networks: A Hybrid KNN-Decision Tree Model for Anomaly Detection and False Alarm Reduction

Anushri Narendra Pathak[1,*] , Arvind R. Yadav[2]

[1]Research Scholar, E&C Dept, Parul Institute of Engineering and Technology, FET, Parul University, Vadodara, India

[2]Associate Professor, E&I Engineering Dept, Institute of Technology, Nirma University, Ahmedabad, India

Emails: anushrip21@gmail.com; arvind.yadav.me@gmail.com

Abstract

In applications related to military operations, Wireless Sensor Military Networks (WSMNs) aid a critical function by deploying a distributed group of sensor nodes. Such sensor networks lift the overall effectiveness of military activities by situational alertness and permitting instantaneous decision-making processes. This deployment also rises noteworthy challenges, namely scalability, energy efficiency, and security vulnerabilities. Ensuring the accessibility, trustfulness and confidentiality of the data sensed by sensor nodes is prime important challenge. It could lead to disastrous consequences on the military field. Looking into this shortfall, ongoing research is mainly targeted at obtaining advanced solutions to such challenges, such as secure and energy-efficient routing algorithms. However, one of the considerable challenges in WSNs is anomaly detection and the existence of false alarms. This can affect the dependability and effectiveness of the system. The ongoing research in this field focuses on exploring the condition of WSMN, mainly their applications, challenges, and future directions. Authors propose an adaptive and hybrid Machine Learning (ML) approach to reduce false alarms and anomaly detection along considering mutual authentication system. ML approaches offer reliable solutions by improving the data classification accuracy and detection of anomalies. These algorithms have better capability to distinguish between normal and abnormal events, which ultimately reduces false triggers. The authors propose a hybrid approach of k-Nearest Neighbors (KNN) and Decision Tree (DT), which results in a powerful method for improved classification accurateness and robustness in WSN. The effectiveness of KNN in local decision-making and better clear interpretability of Decision Tree to handle feature interactions are combined together in this strategy, to increase overall performance.

Keywords: Wireless Sensor Military Networks (WSMN); False Alarms; Energy-Aware; Machine learning; Decision-Tree; K-nearest neighbor

1. Introduction

WSMNs are very important part in current military operations, such as real-time monitoring of environments, potential fears, herd movements, etc. These networks contains a number of sensor nodes that detect and transfer data, providing leaders with awareness of critical situation. Nevertheless, a noteworthy challenge met by WSMNs is the existence of false alarms and anomaly detection [1, 2]. False alarms may result in waste of resources, misallocation of heads, and detachment from genuine threats. This may result ultimately in risking mission success. In military, timely and accurate information is paramount, thus, restraining false alarms is crucial for preserving effectiveness of operation and safety insurance. Anomaly detection is a vibrant process that contains pattern recognition that diverges from expected behaviour. In WSMNs, anomalies indicate events that are critical, such as equipment malfunctions, enemy movements, even they can arise from environmental factors such as communication noise [3,4].

Therefore, distinguishing ability between false alarms and genuine anomalies is of prime importance. The ML algorithms can be used for anomaly identification with larger accuracy by analysing historical data to learn normal

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

behaviour patterns [5]. Techniques that use data fusion improve detection consistency and decrease false positives. Thus, this improves operational efficiency and strengthens decision-making processes in high-stakes military environments.

Anomaly Detection: Various ML algorithms can be trained on past data. These algorithms are SVM- Support Vector Machines, NN-Neural Networks, DT- Decision Trees, etc. This is required for identifying of patterns related to genuine threats. These models can disseminate to changes in the environment resulting in increasing their analytical accuracy.

Feature Selection: Effective techniques used for feature selection help in identifying most related characteristics that have precise event classification. Using techniques such as RFE - Recursive Feature Elimination, PCA- Principal Component Analysis, the data dimensionality can be reduced, resulting in increased computational efficiency.

Ensemble Learning: Merging various ML- Machine Learning models over ensemble techniques can result into better values as compared to individual models. These tactics exploit on the strengths of several algorithms, thus increasing robustness and overall resulting in reduced false alarms.

Adaptive Thresholding: As an alternative of using fixed thresholds required for event detection, real-time data analysis based adaptive thresholds can be useful to adjust dynamic sensitivity levels. ML models can study optimal thresholds, which are content-based data patterns, resulting in minimized false alarms.

1.1. Issues and Challenges

There are different challenges in WSMN. They can be named as anamoly discovery, network security, minimizing false alarm, etc. Guaranteeing of security in WSN is possible by recognition and reduction of malicious activities in the network [6]. The unkonwn attacks are not easy to detect by old-style attacker detection schemes using predefined signatures or thresholds [7]. WSMN's security and anomaly intrusion detection performance can be improved using promising ML tactics [1]. ML algorithms training is done using historical data to identify anomalies or patterns telling of potential intrusions [3], [8]. This approach can develop an intelligent security system [9]. Such structures can examine vast dataset; abnormal patterns can be identified, and discriminate between normal and malicious behaviours [10]. Different ML algorithms, namely random forests, decision trees, neural networks, etc. can extract valuable perceptions from intricate WSMN datasets, resulting in improved accuracy of invasion detection methodologies [1], [10].

Existing surveys in WSMN detection of anomaly have many challenges. Scarce scalability is an important worry, as WSNs often involve a greater number of sensor nodes, due to which computational difficulty increases [11]. Higher rates of false positive and lacking flexibility for evolving attack methods are also some of the challenges that needs a solution [12], [13] & [14]. Additionally, the uneven data distribution in WSN , categorised by a majority of normal instances and a lack of invasion cases, results in obstacles of getting precise detection outcomes. The proposed method represents a prominent improvement in enhancing security and false alarms reduction for WSNs. This strategy is vital in overcoming the inadequacies of current processes and bringing an ideal solution for anomaly detection in WSNs.

2. Related Work

The survey mainly emphases on detection of false alarm as well as fault tolerance in case of sensors that are self-organized and use some authentication methods. In WSMNs, getting security as well as lesser overhead is a giant challenge. For this purpose, many authentication methodologies have been established. Considering the constraints of sensors such as scarce storage, lesser processing power, and communication capabilities, it is needed to produce an effective verification system.

Authors present a self-organizing, resilient system for detection of wildfires in maintenance areas in [15]. The model uses software componentization and diversity techniques (N-Version programming) to create self-adaptive low-power applications and develop fault-tolerant applications. The simulation results show that the model can effectively detect potential failures in the fire detection decision trees and reduce false alarms, leading to energy savings and increased network lifetime. Such model can also be applied to other applications like flood monitoring, pollution monitoring, and smart city applications.

Authors presented ML techniques, which can be used to reduce false alarm rates in wireless intrusion detection systems [16]. Wireless intrusion detection systems can significantly lower their false alarm rates by implementing machine-learning algorithms. The study suggests a two-stage filtration procedure and feature reduction followed by a machine learning-based false alarm filter. The paper used AWID dataset, which is more relevant for wireless networks compared to other commonly used datasets. Dependence on human analysts for training or rule development is observed.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

An Online Locally Weighted Projection Regression (OLWPR) algorithm, having an error rate of only 16% yet a detection rate of 86% has been discussed in [17]. One vital requirement for WSN is low computational complexity, which is owned by the LWPR approach used here. WSNs are prone to contain irrelevant yet redundant data, which should be addressed through methods like online dimensionality reduction. The OLWPR algorithm proposed in the paper has an 86% detection rate and 16% error rate. This suggests there is room for improvement in anomaly detection accuracy.

The authors discussed about network anomaly detection in WSN having the potential of graph-based deep learning [18]. It describes the key features and security fears of WSN. It reviews current studies on network anomaly detection using graph-based deep learning techniques. The limitations of the sensors make them vulnerable to malicious activities. Resolving of challenges is necessary which can be done by using a anomaly detection technique, with suitable datasets for validation, and intricate machine learning models.

The authors presented reduction of false alarms and improvement in target detection accuracy for WSN using machine learning (ML) algorithm in [19]. The efficiency of ML algorithms, including K-Nearest Neighbours, Random Forest, Logistic Regression, and Decision Trees, in dropping false alarms and increasing target identification in WSNs was compared. In real-world applications, it can be seen that incorporating environmental elements, improves target identification accuracy. The results showed how well ML techniques work to reduce false alarms in WSNs. However, only few ML algorithms were examined in this study. In addition, paper only considered weather as an environmental factor, and no other potential environmental factors that could affect sensor readings and target detection.

ML techniques for detecting false data attacks in smart grid systems was surveyed and presented in [20]. The paper discusses the security threats postured by false data attacks in smart grid systems. It also states the importance of effective detection methods. It provides an overview of existing machine learning-based detection methods in areas such as non-technical losses, load forecasting, state estimation, etc.. Existing ML-based detection techniques may be weak in case of adversarial attacks. A collaborative, decentralized detection background could be important future research. It requires techniques for preserving confidentiality in false data attack detection. These deficiencies in the ML-based detection techniques need to be addressed in future research.

A mutual authentication mechanism that is energy-efficient and secure for military wireless sensor networks was described in [21]. Authors suggested solution using energy-aware mutual authentication [21] to enable safe communication between nodes. For prevention of forward and backward secrecy attacks and to generate robust node authentication system, the residual energy of each sensor node was computed. It protects against node analysis-based attacks and prevents compromised nodes from detecting the digital signature from base station. This decreased the communication's computational and memory overhead. In addition, this results in alive sensor nodes with low energy consumption.

The WSMN networks are built with a horizontal design and numerous hops for communication, and the base station is located close to the base camp. This mechanism's structure and method are based on the creation and distribution of digital signatures via a public key cryptography system, which guards against various network layer assaults like wormholes, Sybils, and denial of service. Furthermore, it is imperative to ensure data integrity with respect to both forward and backward confidentiality. The suggested mutual mechanism is mainly considered and used by adversarial sensor nodes in WSMNs to protect communication and counter replay attacks with limited resources. Table 1 presents the summary of the selected significant literatures limitations.

Table 1: Limitations of selected literatures.

Sr. No.	**Authors**	**Limitation**
1	Felipe Taliar Giuntini, Delano Medeiros Beder [15]	Failed to adopt and explore other data mining and AI models like Bayesian networks and neural networks for the fire detection application, leads no grantee the accurate alarms and fault tolerance
2	I. Gethzi Ahila Poornima, B. Paramasivan [17]	The OLWPR algorithm proposed in the paper has an 86% detection rate and 16% error rate, suggesting there is stillroom for improvement in the accuracy of anomaly detection. No assurance of energy ware and security.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

3	Takwa Allaoui, M. Jeridi, T. Ezzedine [19]	The study only compared a limited set of machine learning algorithms, and did not explore other potential algorithms that could further improve performance. The study only considered weather as an environmental factor, and did not explore the impact of other potential environmental factors that could affect sensor readings and target detection
4	R. Leppänen, T. Hämäläinen. [18]	The open and immature nature of wireless sensor networks, as well as the limitations of the sensor nodes themselves, make them vulnerable to malicious activities. The potential and feasibility of using graph-based deep learning for detecting anomalies in wireless sensor networks requires further research and development.
5	Sudaroli Vijayakumar D, Ganapathy S. [16]	Lack of accuracy in some proposed techniques. Inability to detect recent unknown attacks. Unsuitability of current machine learning algorithms for real-time use without protocol changes
6	Lei Cui, Youyang Qu, Longxiang Gao, Gang Xie, Shui Yu. [20]	A collaborative and decentralized detection framework could be an important future research direction. Techniques for preserving privacy in false data attack detection are needed. There are deficiencies in the current ML-based detection mechanisms that need to be addressed through future research.
7	Zhang, Rajan Shankaran, Junqi, Mehmet A. Orgun, Abdul Sattar, and Vijay Varadharajan. [25]	False alarm detection accuracy needs to improve for more unknown events.

Conversely, anomaly detection assumes that any intrusive activity is inherently abnormal. This implies that if we were able to create a typical activity profile for a system, we might theoretically mark as infiltration attempts. Each of these methods have their own pros and cons. As pattern matching do not have signatures, the earlier is unable to detect novel assaults, but it has higher accuracy in detecting recognized attacks.

Few of above protocols can classify new threats, but as it is difficult to create realistic behaviour profiles for protected systems, they have a high false alarm rate. The authors here have created a hybrid algorithm for improving detection accuracy for both known as well as unknown threats, maintaining a low false alarm rate.

3. Proposed System

This model deployed on self-organized sensor with a mutual authentication methodology for military WSNs that is both energy-aware and fault-tolerance communications. In order to provide secure communication between nodes, the proposed system is made to have a mutual authentication coordinator and collaborator along with energy-aware and self-organized network. Each sensor node's residual energy was calculated during the authentication process, for the purpose of energy-aware and authenticated coordinator. The proposed system aims to provide strong node authentication technique. The coordinator performs the tasks of sensor data analysis and filtering the uneven or unknow alarms from the network. The decision tree is used to generate event information once the sensed data has been processed. The rules produced by the decision tree are used to decide node information. To get the result, this event data is fed into the KNN together with the initial set of attributes. The main concept here is to examine whether the decision tree's events will enhance the KNN's performance.

The proposed system of authentication significantly decreased false alarms and attacker communication. This results in longer-lasting sensor nodes that should use less energy. The proposed mechanism and its structure are based on digital signature generation and distribution through a public key cryptography system, which protects against all kinds of network layer attacks such as denial of service, Sybil, and wormhole. In WSMNs, the proposed mutual system primarily aims to achieve the self-organized coordinator taken into consideration and utilized by hostile sensor nodes to safeguard communication and defend against replay attacks with constrained resources. The system consists of the following phases: (1) Dataset preparation (2) coordinator election and configuration; (3) Hybrid Classifier (4) mutual authentication and false alarm detection.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

A. Datasets Preparation

Analysis and feature extraction are crucial steps in the creation of a dataset. Presenting the information that is required to make the communications comprehensible and to enable the dataset to be utilized for automatic classification or prediction techniques is the goal. A large number of datasets for anomaly detection systems with a variety of traffic kinds and attack scenarios were found in the literature. These datasets were either constructed in a real world or simulated environment, or they were combined with existing datasets. The evaluation of the networks which are based on intrusion detection system is most commonly carried out with the help of three data sets namely KDD Cup 1999 Data [21], [22], the NSL-KDD dataset [23] and the Darpa 2000 [24]. These datasets are nothing but the improvements over their previous versions. Certain flaws are associated with KDD and it might not exactly represent real networks. There are very few available datasets for IDS networks. The NSL-KDD dataset overcomes the fundamental issues found in the KDD Cup 1999. The KDD Cup 1999 Data was simulated using a common LAN network used by the US air forces and has a significant number of simple connections with 41 features and 24 types of attacks. They used DoS, R2L, U2R, and probing techniques to gather TCP data over the course of nine weeks.

B. Common network attack's dataset

Denial of Service Attack (DoS): Such attack occurs when an attacker denies access of computer to the authorized users or makes a memory/computing resource too busy to handle requests.

Network Attack (PROBE): An attempt to obtain data about a computer network with the apparent goal of getting over its security measures is known as a network attack (PROBE). The four categories and the attacks that corresponded to each category are displayed in Table 2. The details of Attackers are given in Table 3.

Table 2: Dataset Model and their details

Attacks and Event dataset	**Classes or Attributes**
DOS attack dataset	Pod, Neptune, back, smurf, land, teardrop
Network Attacks (PROBE)	portsweep, nmap, ipsweep
URL(U2R)	Load_module, buffer_overflow
Temperature	CO, CO2,
Objects	Humans, animals
Vehicle	Car, jeep, bus

Table 3: Attacker Details in Dataset

Attack Model	**Description**	**Attack Type**	**Dataset Used**
Pod Attack	A POD (Ping of Death) attack encompasses sending a compromised and malformed or huge packet to a goal machine, basically a ping packet	DDOS	DARPA
Neptune Attack	Neptune attack is a basically a type of Denial of Service (DoS) attack which uses SYN flood procedures to overpower a server	DDOS	DARPA
Back Attack	A Backdoor attack comprises the invader installing software that lets them to sidestep normal verification mechanisms and gain illegal access to a system	DDOS	DARPA
Smurf Attack	A Smurf Attack is a type of Distributed Denial of Service (DDoS) attack which strengthens the influence of the	DDOS	DARPA

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

	attack by manipulating Internet Control Message Protocol (ICMP) traffic		
Land Attack	A Land Attack comprises sending a particularly crafted packet that has the equal source as well as destination IP addresses, and matching source and destination port numbers, effectually "looping" the packet back to the sender	DDOS	DARPA/KPP CUP 1999
Teardrop Attack	A Teardrop Attack exploits vulnerabilities in the way some operating systems handle fragmented packets. The attacker sends malformed or fragmented packets that the target machine cannot reassemble properly	DDOS	KPP CUP 1999
Port Sweep	A Port Sweep is a kind of network scan where in an attacker tries to recognize open ports on a range of IP addresses. The attacker typically sends packets to a range of ports on a target system and looks for responses indicating which ports are open	PROBE	KPP CUP 1999
Nmap (Network Mapper)	Nmap is one of the most popular open-source tools used for network discovery and security auditing. It allows users to scan large networks to discover hosts, services, and open ports, as well as perform more advanced tasks like OS detection, version detection, and script scanning	PROBE	KPP CUP 1999
IP Sweep	An IP Sweep (also known as an IP Scan) is a technique where an attacker scans a variety of IP addresses to get which ones are active (i.e., which systems are live on the network). This is typically done with tools like Nmap or specialized scanning tools.	PROBE	KPP CUP 1999
Load Modules	A load module refers to a part of code or a database that is loaded into memory to be executed. This can include shared libraries, executables, and even dynamically loaded code during the execution of a program	URL	KPP CUP 1999
Buffer Overflow Attacks	When more data is written to a buffer—a temporary storage space—than it can accommodate, the surplus data overwrites nearby memory regions, resulting in a buffer overflow.	URL	DARPA/KPP CUP 1999

C. Sensor Configuration and Coordinator Elections

The sensor nodes are initially thrown at random into the region of interest, where they awaken to create a network. Every sensor node receives requests from its neighbours within its Transmission Range (TR), and in compliance with that, it extracts the node's digital sign. Every node needs to cross-reference digital sign information with neighbouring node storage within TR in order to authenticate inbound transmission. In this configuration, every registered node receives a unique digital signature from the base station, and it is fashioned in compliance with military operations specifications. Each sensor node in the military network can be authenticated using a digital signature. Before the registration step, every node has both its public and private keys, along with the sink node's public key. The training set of fault-tolerance is already loaded in nodes. Meanwhile all cluster nodes are suitable for coordinator; the technique first receives data as input to all group nodes that have predicted the same event messages. The node nearest to the base-station will be coordinator. If there are an equal number of hops, the node with the highest energy level will win. The coordinator of the assessment receives data transmitted by the cooperating nodes. Better network energy balance may be ensured in this way since the election takes into account the fact that coordinating activities demand more energy. As a result, only few nodes will be chosen as coordinator more than once.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

D. Proposed Hybrid Algorithm

The training data D has the following attributes: {A1, A2, … , An}, and each attribute A1 contains the following attribute values: {Ai1, Ai2, … , Aih}. The training data D is given as {t1, … , tn}, where ti = {ti1, … , tih}. Both continuous and discrete attribute values are possible. Additionally, training data D have classes C = {C1, C2, … , Cm}. Every sample in training data D belongs to a specific class, Cj. The method first looks for numerous copies of the same example in training data D. If it finds them, it retains just one unique example in training data D (assuming that two instances are similar if all of their attribute values are equal). In order to discretize the continuous attributes in training data D, the approach first identifies each adjacent pair of continuous attribute values that are not assigned to the same class value for that continuous attribute.

In addition, by means of the class label of the k most comparable neighbours, the algorithm predicts the class of the incoming data and ranks the neighbours of the selected node in the training data set. The likeness of each neighbour of X is used to weight the neighbour classes. Cosine similarity, which is defined as follows by equation (1), is a measure of similarity.

$$Sim(X, Dj) = \frac{\Sigma ti\ \varepsilon\ (X \cap Dj) Xi * dij}{||X||2 * ||Dj||2} \quad (1)$$

Where:
'X' being test node which is represented as a vector;
'Dj' being j^{th} training dataset;
'ti' being the attribute which is been shared by X and Dj
'Xi' being the weight of an attribute ti in X
'Dij' being the weight of attribute ti in Dj
'$||X||^2$' being the normalization of X
'$||\ Dj||^2$' being the normalization of Dj

The algorithm then uses these techniques to classify every sample in training data D. The method determines the information gain for each attribute {A1, A2,…,An} in the training data D if any of the training examples are incorrectly classified.

$$Info(D) = -\sum_{i=0}^{n} \frac{freq(C,Dj)}{|D|}\ log2\left(\frac{freq(C,Dj)}{|D|}\right) \quad (2)$$

$$Info(D) = \sum_{i=0}^{n} \frac{|Ti|}{|T|} \inf\ O(Ti) \quad (3)$$

$$Informartion\ Gain\ (Ai) = info(D) - Info(T) \quad (4)$$

The figure 1. shows the Proposed System Flow. The aggregated data is gathered, processed and normalized on it. The proposed hybrid algorithm is applied so as to improve the performance.

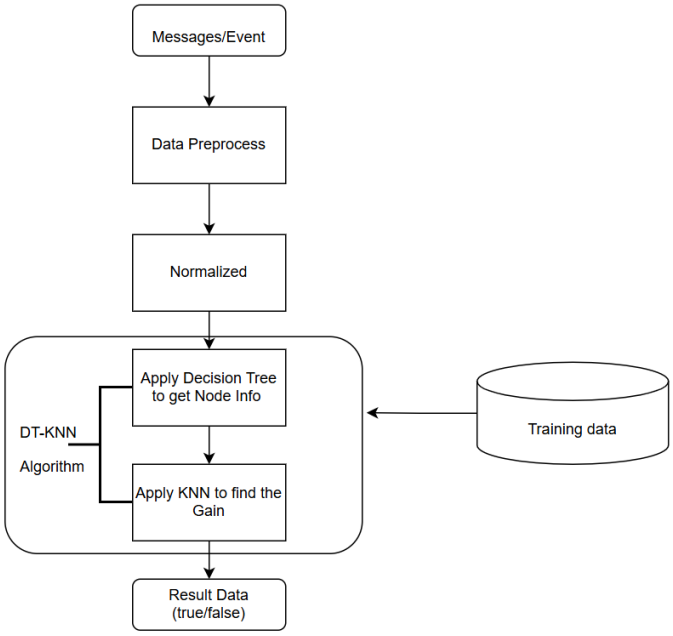

Figure 1. Proposed System Diagram

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

Hybrid DT-KNN Algorithm

Proposed DT-KNN hybrid model combine the strengths of both algorithms to overcome their individual limitations and improve predictive performance. Here's how each algorithm contributes and how they can be integrated:

1. Strengths of Decision Trees (DT):

- Easy interpretability: The decision tree splits the feature space into clear decision regions based on simple decision rules (if-else statements).
- Handles both categorical and numerical data: It can easily deal with both types of variables.
- Non-linear decision boundaries: DT can model non-linear relationships between features.
- Fast training phase: Once the tree is built, predictions are fast because the tree requires just a few decisions to classify a sample.

2. Strengths of K-Nearest Neighbours (KNN):

- Non-parametric: KNN makes no assumptions on the data's distribution.
- Simple and flexible: It works well for complex, non-linear decision boundaries.
- Works well with small datasets: Since KNN is a memory-based algorithm, it can perform well with smaller datasets without the need for explicit training.
- Good performance with noise: KNN can be robust to noise (especially with the right choice of k).

Methodology of Hybrid Model

A DT-KNN hybrid model aims to leverage the strengths of both models, with the idea that a combination can overcome each algorithm's individual weaknesses. The proposed model combines Decision Trees and KNN, as below:

Decision Tree for Feature Selection and KNN for prediction:

- Idea: Use the decision tree to determine important features or to reduce the feature space, and then use KNN for prediction for reduce the false alarms.
- Approach:
  - The decision tree is trained on the full dataset.
  - Important features or a subset of features (e.g., based on feature importance from the decision tree) are selected.
  - KNN is then used to classify the data using only those selected features.
- Advantages:
  - The decision tree helps reduce dimensionality or irrelevant features, which speeds up the KNN algorithm and can lead to better performance by reducing noise.
  - KNN can handle complex decision boundaries and non-linear relationships that a decision tree might miss.

Challenges and Considerations:

- Computational complexity of KNN necessitates calculating the distance between training and test points, which can be computationally costly, particularly when dealing with big datasets. By limiting the search space, the decision tree component of the model can aid in lowering complexity.
- Model tuning: Tuning both the decision tree (depth, pruning, etc.) and KNN (number of neighbours, distance metric, etc.) parameters requires careful consideration and might need cross-validation.
- Interpretability: While decision trees are interpretable, incorporating KNN into the model can make the overall model harder to interpret, particularly if KNN is applied in ways that are more complex.

Mutual authentication and reduced false alarm

The Hybrid framework proposes that the coordinator perform mutual authentication with sensor nodes; and fault-tolerance mechanism. After initialization of network and coordinator selection, each node sends a request message to the base station (BS) in order to request the creation of a digital signature. This contains sensor node IDs, public key, and sensor identity with the coordinator or collaborator as a value hashed MAC nonce encrypted with the

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

base station's public key. In response, if a base station receives encrypted data for each sensor node request, it will use its own private key (PRKBS) to decrypt the data. In the process of mutual authentication, all sensor nodes validate the other node's digital signature through public key of base station (PUKBS) only at selected coordinator (CS). If coordinator (CS) successfully decrypted each sensor collaborator node's identity and digital signature (SignA→N) with BS's public key, then it checked the validity of the signature and verified the residual energy of Sensor nodes. This process ensured that each collaborator was a valid neighbour node to coordinator and stored it in the neighbour table as registered. It is important to note that this occurs in a flexible manner. The binding mechanism is integrated with a machine learning approach to classify tasks to control this dynamic, establish fault tolerance, and provide a correct output. The solution of false alarm detection in WSNs of hybrid decision trees and KNN approach machine learning components: When the sensor node is acting as the coordinator, only then will these hybrid models load and perform the coordination tasks.

4. Results and Discussion

To assess the effectiveness of the proposed security mechanism, we compare the proposed system with self-organized mutual authentication implemented with decision tree classifier and the proposed hybrid algorithm. The authentication key shared with neighbouring sinks is used to validate the authentication ticket. The movement of the node cannot be tracked using the authentication ticket. The neighbouring sinks of the sink that issued the ticket can verify the authenticity of the ticket. We are using an NS-2 simulation military scenario and compared the simulation results with above discussed sensor network protocols. In the simulation, two-ray ground radio propagation is used, the link data rate is 11Mbps, and the sensor nodes have been fitted with 802.11b multi radios. As secure communication uses magnitude more energy for sending or receiving data packets, we only considered the resulting energy cost in the evaluation [25]. Wireless sensor networks usually use default simulation parameters, which are listed in [26-27].

This demonstrates that the proposed hybrid algorithm is successful in identifying false alarm and anomaly detection when it is deployed with coordinator in military network environment. The method used here filter the error probability by measuring the amount of true and false alarms. Every algorithm was trained on the KDD data by means of a 10-fold cross validation test mode for validating its effectiveness. Process use 10-fold cross-validation for testing the algorithms. The test set is always one of the ten subsets, whereas the training set is made up of the remaining k-1 subsets. For each of the ten trials, performance statistics are computed.

This offers a reliable predictor of the classifier's performance on unknown data. For a typical class, hybrid DT-KNN performs better than both KNN and individual decision trees. It outperformed KNN and individual decision tree technique for the Probe and Normal classes. Finally, we have compared with another hybrid algorithm naïve bayes with K nearest neighbour approach [28] for same false alarm detection. In the context of a comparison, it typically showed that the proposed hybrid approach is more optimized with respect of its collective accuracy.

Table 4: Detection Accuracy

Sensed Data	**Classifier Accuracy**					
	Decision Tree		**Hybrid DT-KNN**		**Naïve Bayes + KNN**	
Normal	100%	Average = 97 %	100%	Average = 99 %	99.62%	Average = 97 %
DOS	96.83%		99.11%		95.4%	
PROBE	97.33%		98.56%		97.97%	
URL	94.77%		97.54%		96.33%	
OBJECTS	98.23%		99.01%		97.54%	

Utilizing the hybrid method, we can determine and detect the most of the erroneous events and reduce the false alarm on filtered results. Table 4 shows Detection Accuracy of compared protocols. The same can be observed in pictorial form in Figure 2. In terms of detection percentages, true or false would be the prevailing results in this instance. Because the outcomes of two choice trees are true, the conclusion is true. Because it is the probability average of them employing hybrid approach with true as its conclusion, the probability result is 99.01%. We have added 20% percentage of attacker dataset while driving the detection accuracy.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

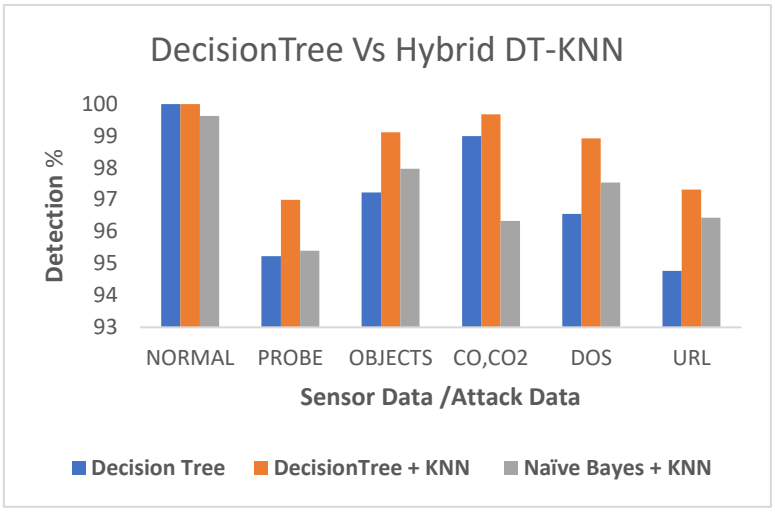

Figure 2. Detection Accuracy

Another important parameter is False Alarm detection by hybrid approach as shown below, 20% false data is inputted to Hybrid approach which primary drive the number false alarm count by each attacker. The Table 5 lists the false alarm count with respect of attackers for existing and proposed Hybrid approach. Figure 3. depicts pictorial representation of the data given in table 3.

Table 5: False Alarm Detection Count

Sensed Data	False Data Count & Accuracy					
	Decision Tree		Hybrid DT-KNN		Naïve Bayes + KNN	
CO, CO2	399	Average = 365.4	489	Average = 464.4	432	Average = 437.2
DOS	388		450		412	
URL	340		481		466	
PROBE	367		435		423	
OBJECTS	333		467		453	

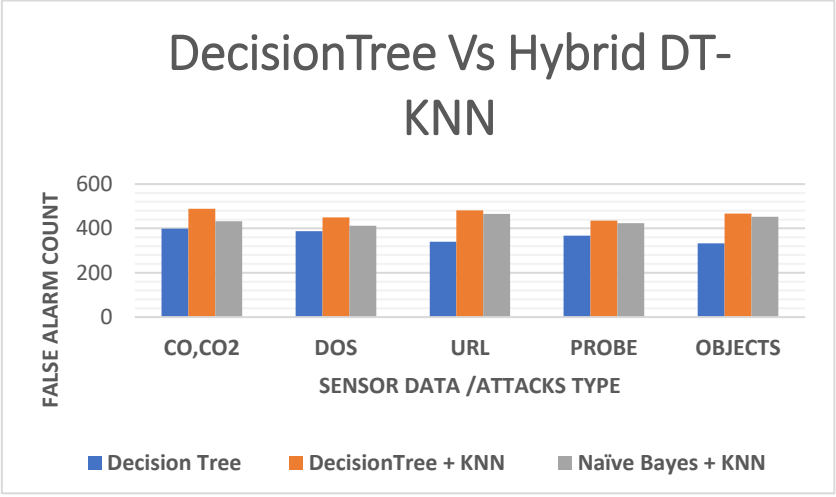

Figure 3. False Alarm Count

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

The accuracy of each algorithm's successfully classified occurrences is shown in the figures above. The KDD test dataset's detection and classification make it clear that the hybrid DT-KNN method has a higher detection accuracy. The comparison of the proposed Hybrid simulation model is done with various parameters like changing number of rounds and changing network load (packet size) with respect of energy, delay and throughput graphs.

A. Simulation Time vs Hybrid Approach

The average amount of energy used when sending the data and aids in estimating how long a wireless sensor network will last. A node's energy usage is examined, taking into account the energy used for event detection, packet transmission, and packet reception. It is an especially crucial metric for assessing the WSMN's success. Figure 4. shows the average energy usage of the suggested method compared to the current one. Simulation time various like 50, 100, 150 and the packet size 512 kbs/sec data rate used for graphs comparison. Table 6 shows Average Remaining Energy for different no. of Rounds, and overall improvement of proposed hybrid protocol.

Table 6: Average Remaining Energy for different no. of Rounds

Number Of Rounds	Average Remaining Energy n Joules		% Observed Improvement over DT
	Decision Tree	Hybrid DT-KNN	
50	4.624987	4.653192	0.61 %
75	4.412997	4.512997	2.27 %
100	4.268334	4.368334	2.34 %
125	4.056279	4.256279	4.93 %
150	3.949161	4.049161	2.53 %
175	3.824945	3.924945	2.61 %
200	3.758128	3.858128	2.66 %

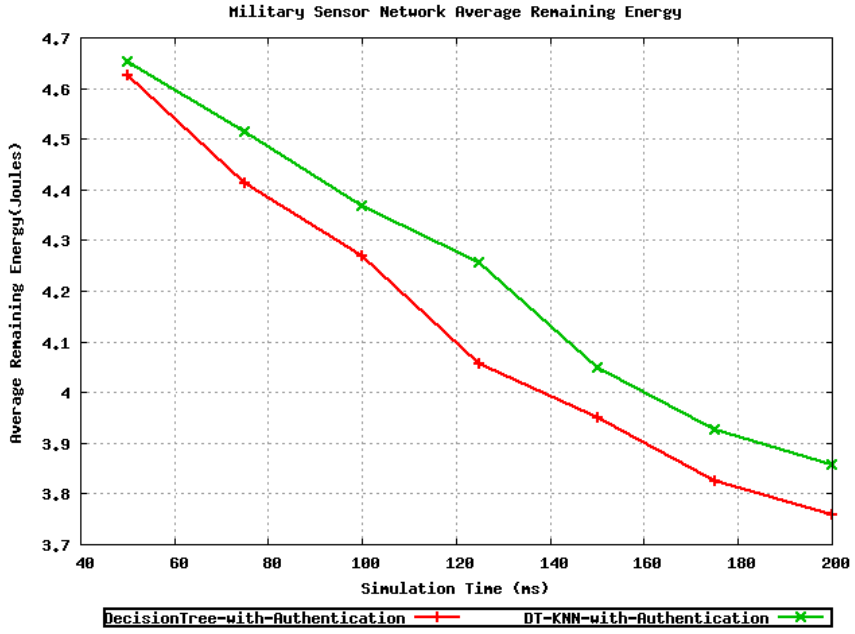

Figure 4. Average Remaining Energy vs Time Graph

Figure 5. shows detected false alarm count by both algorithms. Table 7. shows False Alarm Detection for different no. of Rounds, and overall improvement of proposed methodology.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

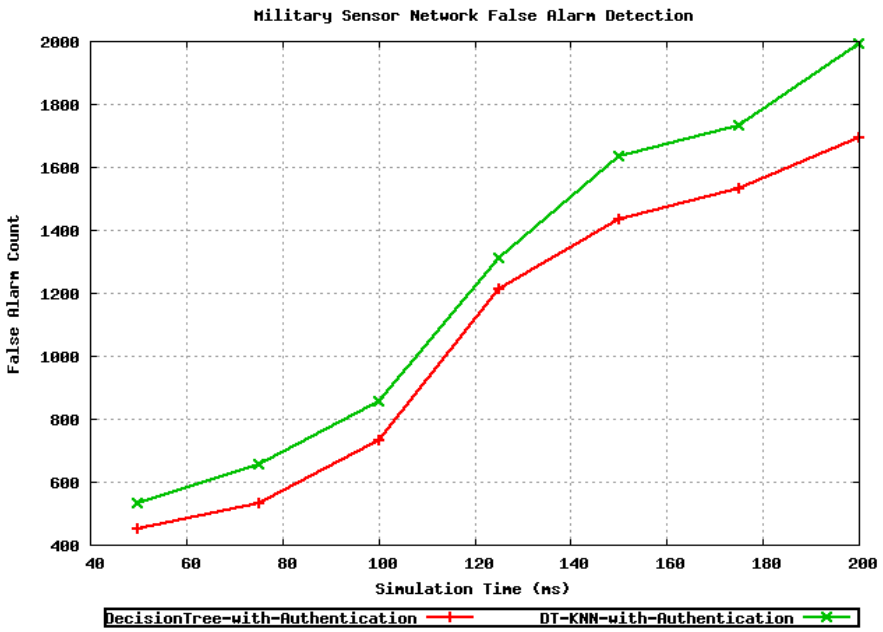

Figure 5. False Alarm Count Vs Time Graph

Table 7: False Alarm Detection for different no. of Rounds

Number Of Rounds	**False Alarm Detection in Counts**		**% Observed Improvement over DT**
	Decision Tree	**Hybrid DT-KNN**	
50	450	532	18.22 %
75	534	654	22.47 %
100	733	854	16.51 %
125	1211	1311	8.26 %
150	1432	1632	13.97 %
175	1533	1733	13.05 %
200	1693	1993	2 %

B. Packet Size Vs Hybrid approach.

The average amount of energy used when sending data and aids in estimating how long a wireless sensor network will last. A node's energy usage is examined, taking into account the energy used for event detection, packet transmission, and packet reception. Figure 6. shows that the hybrid approach has less usage sensor energy for detection and transferring the packet even it high network load. Table 8 shows Average Remaining Energy for different packet sizes, and overall improvement of the proposed methodology.

Table 8: Average Remaining Energy for different packet sizes

Packet Size Kb/sec	**Average Energy Remaining in Joules**		**% Observed Improvement over DT**
	Decision Tree	**Hybrid DT-KNN**	
64	3.849161	4.293192	11.54
256	3.724945	4.112997	10.42

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

512	3.658128	4.068334	11.21
778	3.449743	3.956279	14.68
1024	3.326545	3.849161	15.71
1280	3.158128	3.724945	17.95
1580	3.023211	3.658128	21.00

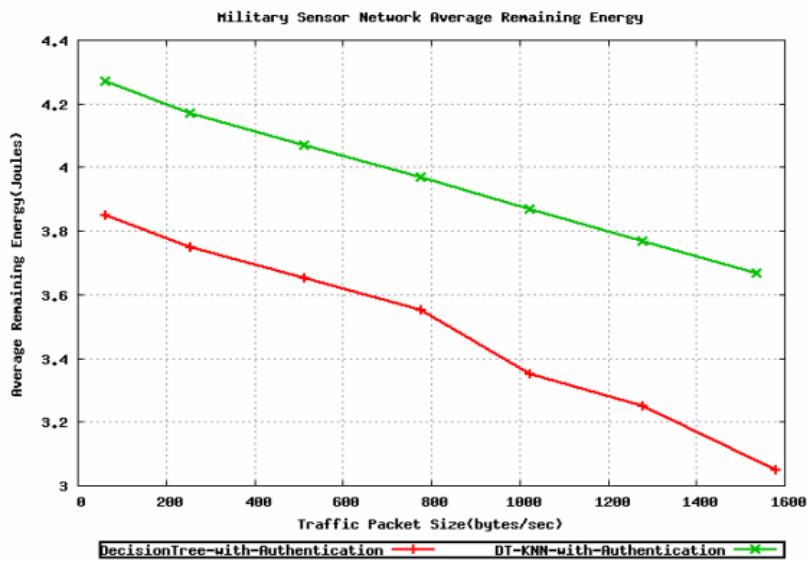

Figure 6. Average Remaining Energy vs Packet Size

Figure 7.shows the false alarm count while increasing the packet size of the proposed method compared to the current one. From the graph the hybrid approach doing well performance while increase the network load. The packet size various from 256 kbs/sec to 1580 kbs/sec data rate used for graphs comparison. Table 9 shows False Alarm Detection for different packet sizes, and overall improvement of the proposed methodology.

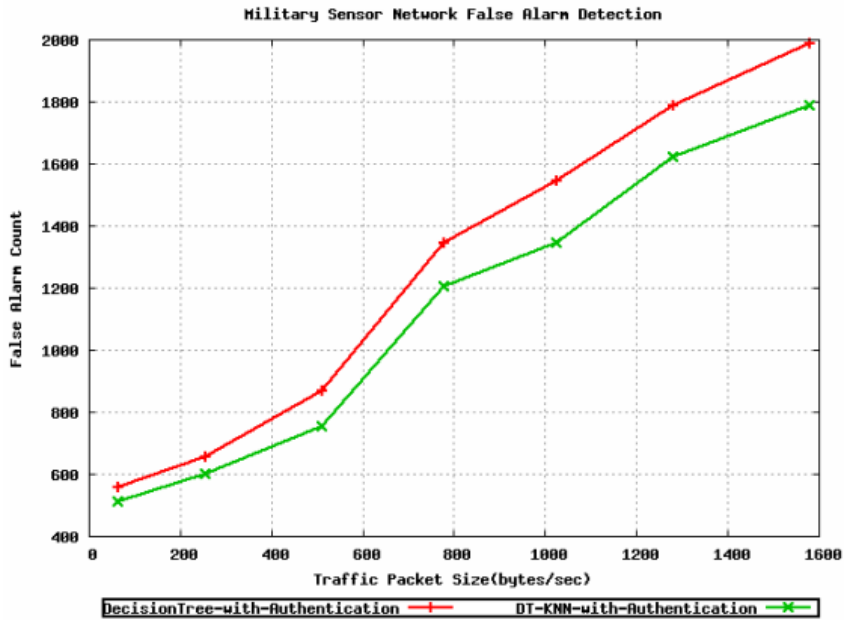

Figure 7. False Alarm Count Vs Packet Size

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

Table 9: False Alarm Detection for different packet size

Packet Size Kb/sec	False Alarm Detection in Counts		% Observed Improvement over DT
	Decision Tree	Hybrid DT-KNN	
64	512	556	8.59
256	598	656	9.70
512	754	866	14.85
778	1203	1346	11.89
1024	1365	1533	12.31
1280	1645	1788	8.69
1580	1764	1976	12.02

The Figure 8. shows Throughput of Hybrid approach and DT approach. Each approach performs well while transferring the data in terms of security, yet the proposed hybrid algorithm outperformed when compared with DT. Table 10 shows Packet Throughput for different packet sizes, and overall improvement of the proposed methodology.

Table 10: Packet Throughput for different packet size

Packet Size Kb/sec	**Packet Throughput**		**% Observed Improvement over DT**
	Decision Tree	**Hybrid DT-KNN**	
64	9516	9566	0.53
256	9646	9886	2.49
512	9832	9987	1.58
778	9994	10112	1.18
1024	10523	10789	2.53
1280	10919	11219	2.75
1580	11765	11865	0.85

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

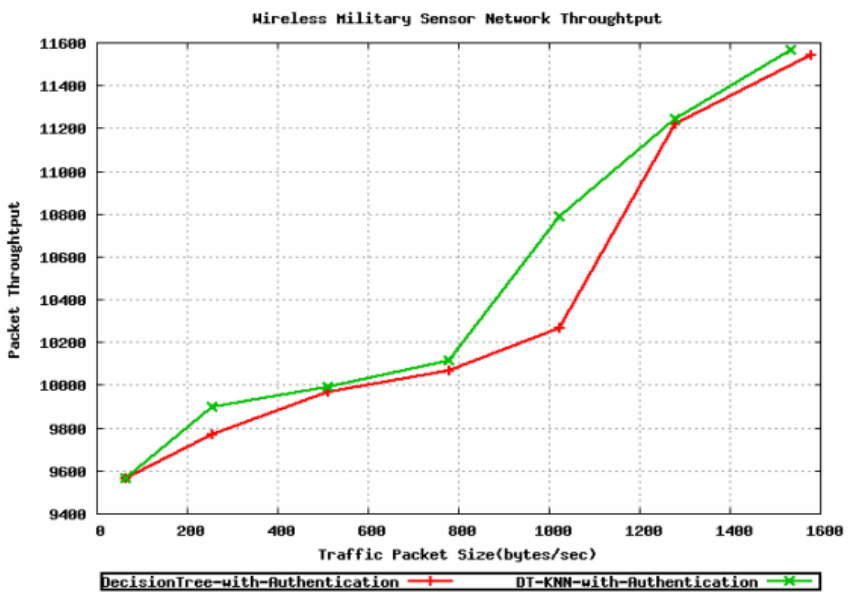

Figure 8. Packet Throughput vs Packet Size

6. Conclusion

False alarms continue to be a major problem in wireless military sensor networks and attacker detection, which limits their development. Building a false alarm filter seems like a good way to cut down on false alarms. In order to improve detection accuracy, lower false alarm rates, and simplify computations, we described in detail how to reduce false alarms using machine-learning techniques like the hybrid DT-KNN approach and preprocessing techniques like filter and aggregation approaches.

The hybrid model's detection accuracy level was determined by the comparison analysis. Following the study, it is suggested to use the proposed hybrid model for improving detection accuracy by lowering the computation complexity and false alarm rate. In order to defend against attackers and lower false alarms, it was finally implemented with intelligent mutually authenticated coordinator-based sensors. With a 99% accuracy rate, proposed method analysis offers reliability and correctness of the proposed work. In addition, when examined for false alarm detection capability, the proposed algorithm provides 27% improvement over Decision Tree algorithm and 6 % improvement over Naïve Bayes + KNN. The same strategy was implied and verified for different number of rounds and packet size, for which improved results are being obtained.

To get the best feature set and the utmost prediction, the majority of the top machine learning algorithms are also researched. In order to improve forecast accuracy for upcoming assaults, it recommended developing new machine learning algorithms using fuzzy rules, soft computing approaches, and more sophisticated military sensors.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] M. P. Đurišić, Z. Tafa, G. Dimić, and V. Milutinović, "A survey of military applications of wireless sensor networks," in *2012 Mediterranean Conference on Embedded Computing (MECO)*, 2012, pp. 196-199.

[2] G. Singh, "Security attacks and defense mechanisms in wireless sensor network: A survey," *International Journal of Innovative Science, Engineering & Technology*, vol. 3, no. 4, pp. 129-136, Apr. 2016.

[3] F. T. Giuntini, D. M. Beder, and J. Ueyama, "Exploiting self-organization and fault tolerance in wireless sensor networks: A case study on wildfire detection application," *International Journal of Distributed Sensor Networks*, vol. 13, no. 4, pp. 1550147717704120, Apr. 2017.

[4] S. Pragadeswaran, S. Madhumitha, and S. Gopinath, "Certain investigation on military applications of wireless sensor network," *International Journal of Advanced Research in Science, Communication and Technology*, vol. 3, no. 1, pp. 14-19, Mar. 2021.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

[5] M. A. Talukder, M. M. Islam, M. A. Uddin, A. Akhter, K. F. Hasan, and M. A. Moni, "Machine learning-based lung and colon cancer detection using deep feature extraction and ensemble learning," *Expert Systems with Applications*, vol. 205, p. 117695, Nov. 2022.

[6] M. A. Talukder, K. F. Hasan, M. M. Islam, M. A. Uddin, A. Akhter, M. A. Yousuf, F. Alharbi, and M. A. Moni, "A dependable hybrid machine learning model for network intrusion detection," *Journal of Information Security and Applications*, vol. 72, p. 103405, Feb. 2023.

[7] A. Heidari and M. A. Jabraeil Jamali, "Internet of Things intrusion detection systems: a comprehensive review and future directions," *Cluster Computing*, vol. 26, no. 6, pp. 3753-3780, Dec. 2023.

[8] A. Sezgin and A. Boyacı, "AID4I: An Intrusion Detection Framework for Industrial Internet of Things Using Automated Machine Learning," *Computers, Materials & Continua*, vol. 76, no. 2, 2023.

[9] T. M. Ghazal, "Data Fusion-based machine learning architecture for intrusion detection," *Computers, Materials & Continua*, vol. 70, no. 2, pp. 3399-3413, 2022.

[10] S. Ifzarne, H. Tabbaa, I. Hafidi, and N. Lamghari, "Anomaly detection using machine learning techniques in wireless sensor networks," *Journal of Physics: Conference Series*, vol. 1743, p. 012021, 2021.

[11] S. Sharmin, I. Ahmedy, and R. Md Noor, "An energy-efficient data aggregation clustering algorithm for wireless sensor networks using hybrid PSO," *Energies*, vol. 16, no. 5, p. 2487, Mar. 2023.

[12] X. Tan, S. Su, Z. Huang, X. Guo, Z. Zuo, X. Sun, and L. Li, "Wireless sensor networks intrusion detection based on SMOTE and the random forest algorithm," *Sensors*, vol. 19, no. 1, p. 203, Jan. 2019.

[13] N. M. Alruhaily and D. M. Ibrahim, "A multi-layer machine learning-based intrusion detection system for wireless sensor networks," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 4, pp. 281-288, 2021.

[14] R. Rathore and M. Hussain, "Simple, secure, efficient, lightweight and token based protocol for mutual authentication in wireless sensor networks," in *Emerging Research in Computing, Information, Communication and Applications: ERCICA 2015*, vol. 1, 2015, pp. 451-462.

[15] F. T. Giuntini, D. M. Beder, and J. Ueyama, "Exploiting self-organization and fault tolerance in wireless sensor networks: A case study on wildfire detection application," *International Journal of Distributed Sensor Networks*, vol. 13, no. 4, pp. 1550147717704120, Apr. 2017.

[16] D. Sudaroli Vijayakumar and S. Ganapathy, "Machine learning approach to combat false alarms in wireless intrusion detection system," *Computer and Information Science*, vol. 11, no. 3, pp. 67-81, 2018.

[17] I. G. Poornima and B. Paramasivan, "Anomaly detection in wireless sensor network using machine learning algorithm," *Computer Communications*, vol. 151, pp. 331-337, Feb. 2020.

[18] R. F. Leppänen and T. Hämäläinen, "Network anomaly detection in wireless sensor networks: A review," in *Internet of Things, Smart Spaces, and Next Generation Networks and Systems: 19th International Conference, NEW2AN 2019, and 12th Conference, ruSMART 2019*, St. Petersburg, Russia, Aug. 26–28, 2019, pp. 196-207.

[19] T. Allaoui, M. H. Jeridi, and T. Ezzedine, "False alarm reduction in WSN surveillance application through ML techniques," in *2023 International Wireless Communications and Mobile Computing (IWCMC)*, 2023, pp. 996-1001.

[20] L. Cui, Y. Qu, L. Gao, G. Xie, and S. Yu, "Detecting false data attacks using machine learning techniques in smart grid: A survey," *Journal of Network and Computer Applications*, vol. 170, p. 102808, Nov. 2020.

[21] S. Rosset and A. Inger, "KDD-Cup 99: Knowledge Discovery in a Charitable Organization's Donor Database," *SIGKDD Explorations*, vol. 1, pp. 85-90, 2000.

[22] N. Paulauskas and J. Auskalnis, "Analysis of data pre-processing influence on intrusion detection using NSL-KDD dataset," in *2017 Open Conference of Electrical, Electronic and Information Sciences (eStream)*, 2017, pp. 1-5.

[23] M. Tavallaee, E. Bagheri, W. Lu, and A. A. Ghorbani, "A detailed analysis of the KDD CUP 99 data set," in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, 2009, pp. 1-6.

[24] DARPA Intrusion Detection Evaluation. Available online: https://archive.ll.mit.edu/ideval/data/2000data.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025

[25] J. Zhang, R. Shankaran, M. A. Orgun, A. Sattar, and V. Varadharajan, "A dynamic authentication scheme for hierarchical wireless sensor networks," in *International Conference on Mobile and Ubiquitous Systems: Computing, Networking, and Services*, 2010, pp. 186-197.

[26] L. Cheng, J. Niu, J. Cao, S. K. Das, and Y. Gu, "QoS aware geographic opportunistic routing in wireless sensor networks," *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 7, pp. 1864-1875, Jul. 2014.

[27] D. R. Raymond and S. F. Midkiff, "Denial-of-service in wireless sensor networks: Attacks and defenses," *IEEE Pervasive Computing*, vol. 7, no. 1, pp. 74-81, Jan. 2008.

[28] H. Om and A. Kundu, "A hybrid system for reducing the false alarm rate of anomaly intrusion detection system," in *2012 1st International Conference on Recent Advances in Information Technology (RAIT)*, 2012, pp. 131-136.

DOI: https://doi.org/10.54216/FPA.200109
Received: January 01, 2025 Revised: March 03, 2025 Accepted: April 04, 2025