Hindawi
Mobile Information Systems
Volume 2023, Article ID 7690322, 11 pages
https://doi.org/10.1155/2023/7690322

Research Article

Cognitive Lightweight Logistic Regression-Based IDS for IoT-Enabled FANET to Detect Cyberattacks

Khaista Rahman,[1] **Muhammad Adnan Aziz**,[2] **Nighat Usman**,[3] **Tayybah Kiren**,[4] **Tanweer Ahmad Cheema**,[1] **Hina Shoukat**,[5] **Tarandeep Kaur Bhatia**,[6] **Asrin Abdollahi**,[7] **and Ahthasham Sajid**[8]

[1]*Department of Electronic Engineering, School of Engineering and Applied Sciences, Isra University, Islamabad, Pakistan*
[2]*FoIT & CS, University of Central Punjab, Lahore, Pakistan*
[3]*Department of Computer Sciences, Bahria University Islamabad Campus, Islamabad, Pakistan*
[4]*Department of Computer Science (RCET), University of Engineering and Technology, Lahore, Pakistan*
[5]*Department of Computer Science, COMSATS University Islamabad, Attock Campus, Islamabad, Pakistan*
[6]*School of Computer Science, University of Petroleum & Energy Studies (UPES), Dehradun, Uttarakhand, India*
[7]*Department of Electrical Engineering, University of Kurdistan, Sanandaj, Iran*
[8]*Department of Computer Science, FICT, BUITEMS, Quetta, Pakistan*

Correspondence should be addressed to Asrin Abdollahi; a.abdollahi@eng.uok.ac.ir

Received 12 October 2022; Revised 22 October 2022; Accepted 25 November 2022; Published 29 April 2023

Academic Editor: Junaid Shuja

In recent few years, flying ad hoc networks are utilized more for interconnectivity. In the topological scenario of FANETs, IoT nodes are available on ground where UAVs collect information. Due to high mobility patterns of UAVs cause disruption where intruders easily deploy cyberattacks like DoS/DDoS. Flying ad hoc networks use to have UAVs, satellite, and base station in the physical structure. IoT-based UAV networks are having many applications which include agriculture, rescue operations, tracking, and surveillance. However, DoS/DDoS attacks disturb the behaviour of entire FANET which lead to unbalance energy, end-to-end delay, and packet loss. This research study is focused about the detail study of machine learning-based IDS. Also, cognitive lightweight-LR approach is modeled using UNSW-NB 15 dataset. IoT-based UAV network is introduced using machine learning to detect possible security attacks. The queuing and data traffic model is utilized to implement DT, RF, XGBoost, AdaBoost, Bagging and logistic regression in the environment of IoT-based UAV network. Logistic regression is the proposed approach which is used to estimate statistical possibility. Overall, experimentation is based on binomial distribution. There exists linear association approach in logistic regression. In comparison with other techniques, logistic regression behaviour is lightweight and low cost. The simulation results presents logistic regression better results in contrast with other techniques. Also, high accuracy is balanced well in optimal way.

1. Introduction

Integration of 5G wireless networks with FANETs is a new concept which uses to improve coverage and reduce delay [1–3]. Mobile ad hoc network is considered the primary idea where VANET and FANET are emerged. UAV swarms or group collectively make FANETs [4]. There can be either signal or multi-UAVs system. Initially, UAVs are only utilized to collect data from ground IoT nodes [5]. But, nowadays, aerial vehicles have changed the dynamics of every human which include smart farming using UAVs, rescue operations, border surveillance, and many more.

In comparison with other traditional fields, FANETs are very much cost low and can be deployed everywhere. The high mobility patterns of UAVs limit energy level in entire network. Due to wireless connectivity in FANETs, internet of things plays an important role. Although, there exist two ways of communication which consist of a2a (air-to-air) and a2g (air-to-ground) [6]. Recently, Zigbee (IEEE 802.15.4) is introduced in FANETs for secure and long-range communication. Mobile UAV pattern effects quality of service (QoS) in the field of IoT-based FANETs. In the conventional UAV network, there exist satellite, ground base station, and UAVs [7].

FANET network needs to be secured from cyberattacks which reduce connectivity in between nodes and interrupt communication. False data attack is one of the dangerous threats during remote patient surgery or operation [8]. However, DoS/DDoS security attacks can be easily detected with the help of the intrusion detection system. Various research studies formulate that identification of cyberattacks in FANETs is considered a major problem [9]. Intruder/attacker UAVs can be used to steal data and jam potential links [10, 11]. Therefore, a proposed system model will consists of detecting ongoing cyberattacks like DoS/DDoS and ping of death which is referred to as dynamic-IDS. This research study will only expand to simulate detection of attacks in FANETs. Furthermore, topological arrangements of FANETs are shown in Figure 1. The main points of the research paper are as follows:

(i) Machine learning algorithms such as DT, RF, XGBoost, AdaBoost, Bagging and logistic regression are utilized

(ii) UNSW-NB 15 dataset is used for training and testing data

(iii) Cognitive lightweight-LR approach is proposed to detect attacks

(iv) Detailed comparative analysis is formulated using machine learning techniques

Major contribution points of this study elaborate the concept of machine learning algorithms which use to detect possible cyberattacks. Comprehensive study is evaluated to understand previous ideas and compare them with the proposed solution. UNSW-NB15 dataset is utilized for experimentation and performance analysis of machine learning classifiers.

Figure 1 illustrates the concept of UAV network using the concept of intruders. When unmanned aerial vehicles tries to collect data from IoT ground nodes at the same time attackers use to deploy fake data packets which leads miss information. Also, FANET network is presented which use to have base station, satellite, and UAVs.

Apart from that machine learning techniques are used in IoT, ad hoc networks, software define networks, and many other fields. Therefore, in machine learning data set is utilized which use to have detailed data for the specific area. Classifiers or algorithms are trained properly to evaluate the performance.

The rest of the article is structured with Section 1 which consists of the study introduction where Section 2 is composed of brief literature having past data about the problem. Similarly, machine learning algorithms in Section 3 and Section 4 represent the proposed model. Section 5 demonstrates simulation results. The theoretical analysis and future direction is discussed in Section 6, which is explained in the conclusion section.

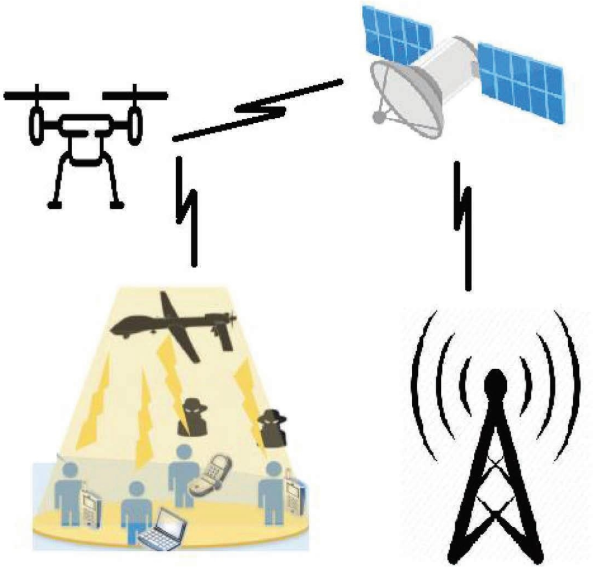

Figure 1: Physical arrangement of UAV network.

2. Related Survey

In the literature section, limitations regarding traditional IDS in other fields are discussed as follows.

Initially, IDS was designed for MANET, VANET, WSN, and IoT networks which use to be vulnerable to cyberthreats such as sinkhole, DoS/DDoS, and PoD. Sometimes, inside the network, attack is initiated which is commonly called sinkhole. While, due to DoS/DDoS security attacks the other neighbor nodes become unavailable for legitimate user. Abdollahi and Fathi implemented a novel IDS for internet of things to identify abnormal data packets. Furthermore, false alarm and missed detection should be reduced which cause issues in network [12].

Real-time IDS can capture abnormal live data packets in contrast with offline. KDD cup 99 data set is commonly utilized in machine learning algorithms to detect damaged caused because of cyberattack [13]. Therefore, real-time IDS are needed for recently emerged technology FANETs.

Identification of attacker through IDS is widely used approach. Therefore, network-IDS usually collect data from network through monitoring traffic. While, different signs of intrusion and alert messages need detection otherwise IoT network level becomes slow down. Deep learning algorithms using KDD cup 99 is simulated through normal, DoS, Probe, R2L, and U2R where high accuracy is examined. FANET is low cost but intrusion can be happen quite easily due to high mobility. Moreover, this study elaborates intelligent intrusion detection framework for UAVs. Authors proposed signature-based IDS for FANETs [14].

Flooding attacks slow down entire process of FANET networks. UAV-IDS-2020 is utilized which use to have unidirectional and bidirectional flow in the data traffic management [15]. Table 1 presents various IDS in UAV networks. In addition, information in Table 1 is mostly about signature-based intrusion detection system. Various areas of studies are conducted to identify cyberattacks. Also, advance datasets are utilized for experimentation of different machine learning techniques.

3. Machine Learning Classifiers/Algorithms

Machine Learning is a term used in the computer science branch that considerably desires to allow computers to "understand" without being instantly programmed [21, 22]. Computers "understand" in machine learning by enhancing their implementation at assignments through "background." In general, "background" usually implies suiting to information; therefore, there is not an exact border among machine learning and statistical techniques [23]. Machine learning techniques have demonstrated significant assurance in furnishing answers to complicated issues [24]. A few of the applications we employ every day from exploring the Internet to recognizing the speech are the instances of enormous strides created in recognizing the assurance of machine learning [25]. Machine learning have two categories: first is supervised learning and second is unsupervised learning. These two categories will wrap all the combination of classification, and techniques of clustering [26]. Supervised learning strategies enclosed combination of various base classifiers; whereas, unsupervised learning strategies enclosed anticipation maximization algorithms as well clustering techniques. In addition, machine learning techniques are used in different field of studies to improve overall performance.

3.1. Decision Tree. Machine learning is the method of learning or dragging unique designs from extensive data sets by applying techniques from artificial intelligence. Category and forecast are the strategies employed to make out essential data categories and indicate a probable trend [27]. The decision tree is an essential category approach in the machine learning classification. It is typically employed in commerce, management, and detection of fraud [28]. As the typical approach of the decision tree, ID3, C4.5, and C5.0 methods have the values of increased organizing rate, powerful learning capability, and straightforward structure. Yet, these methods are also insufficient in a functional application [29]. When utilizing it to categorize, there exists the issue of bending to select features that have more weight and managing features that have fewer weights. Decision trees are amazing techniques to enable anyone to determine the most suitable method of activity [30]. They develop a favorably beneficial arrangement in which one can set choices and investigate the potential consequences of those choices. A decision tree is employed to describe graphically the findings, the possibilities, and the results related to conclusions and occurrences [31].

3.2. Random Forest. Random forest is a unique approach in the field of machine learning that solves many complex issues [32]. Random forest is a mixture of a sequence of tree network classifiers. This approach has numerous useful features and has been significantly employed in the categorization, forecasting, and regression process [33]. Corresponding with the classic approaches random forest has numerous useful integrities; thus, the extent of the application of this unique approach is extremely comprehensive [34]. It is one of the most suitable learning approaches. Generally, this technique is a regression-tree approach that employs bootstrap collection and randomization of forecasters to acquire an increased extent of predictive accurateness [35]. The principal disadvantage of this unique approach is that an enormous number of trees can make the approach slow and inadequate for real-time forecasts. Generally, these approaches are quick to prepare, but a little slow to make forecasts once they are prepared [36].

3.3. Extreme Gradient Boosting. The XGBoost is a brief name for the extreme gradient boosting technique. It is a unique approach that is also known as a tree-based strategy that poses beneath the supervised component of the machine learning domain [37]. Although it can be employed for both categorization as well as regression issues, all of the instructions and illustrations in this technique guide the algorithm's service for categorization only [38]. It is an important and scalable performance of gradient enabling framework. It sustains diverse accurate operations, involving deterioration, categorization, and ranking [39]. In comparison to the regular gradient boosting, XGBoost employs its strategy of creating trees where the score of the similarity and growth choose the most suitable node breaks [40].

3.4. AdaBoost. Boosting algorithm is a famous approach in the machine learning domain to solve the complex problems. AdaBoost is the standard approach in the family of Boosting [41]. This approach has the authority of resisting overfitting. Comprehending the secrets of this sensation is a charming fundamental academic issue. Multiple investigations are dedicated to describing it through statistical theory and margin approach [42]. AdaBoost approach was the preferably suitable boosting algorithm and stayed one of the most widely employed and examined, with applications in multiple domains. Also, this approach can be utilized to facilitate the execution of any algorithm used in machine learning [43]. These are approaches that accomplish precision just beyond random event on a categorization issue. The most appropriate and hence common method employed with AdaBoost are decision trees along with level one [44].

3.5. Bagging Classifier. Bagging is a widely known ensemble building strategy, where an individual classifier in the ensemble is prepared on a separate bootstrap replicate of the training group [45]. The current outcome has demonstrated that bagging can decrease the effect of outliers in training data, particularly if the distant

Table 1: Intrusion detection system for UAV networks.

Reference	Authors	Field of study	Type of IDS	Description
[16]	Bouhamed et al.	UAV network	Signature	Deep reinforcement learning IDS is formulated to detect possible security attacks
[17]	Shrestha et al.	UAV network	Signature	UAV-based IDS is designed using machine learning where decision tree approach given the optimal results in terms of accuracy
[18]	Amouri et al.	Mobile Internet of things (M-IoT)	Signature	Two different mobility models are used which include random way point and Gauss–Markov while designing IDS for mobile IoTs. DDoS and black hole attacks are easily detected with 98% high power velocity
[19]	Khan et al.	Internet of things	Signature	Deep learning-based IDS is implemented for IoT to detect MITM, DDoS, and DoS
[20]	Ghaleb et al.	Vehicular ad hoc networks (VANETs)	Signature	Machine learning algorithms are used to propose IDS for VANETs using NSL-KDD data set

observations are resampled with a more inferior possibility [46]. It is also known as Bootstrap aggregating, which involves having individual models in the ensemble voice with similar significance. To facilitate sample variance, bagging trains every model in the ensemble employing a randomly marked subset of the training group [47]. As an instance, the random forest approach incorporates random decision trees along with bagging to acquire extremely elevated classification precision. Bagging attempts to execute parallel trainees on undersized sample inhabitants and then carries a norm of all the forecasts [48]. Bagging operates by integrating forecasts by voting, every model obtains equivalent significance "Idealized" interpretation: Model several training groups of size *n* and then create a classifier for each training group and connect the classifiers' forecasts [49].

4. Cognitive Lightweight Logistic Regression Approach

Logistic regression approach is employed to estimate the statistical importance of individual separate variable with reference to possibility [50]. It is a strong form of modelling binomial effect. For instance: if the individual is stirring to suffer from cancer or not by carrying weights 0 as well as 1. Decision trees, as well as logistic regression, are extremely famous approaches in the machine learning domain to solve complex issues [51]. Instead of having so many advantages, decision trees tend to have issues handling linear associations among variables as well as logistic regression has problems with relations effects among variables [52, 53]. Therefore, logistic regression is lightweight and cognitive in nature. Due to lightweight behaviour, LR is easy to deploy on the UAV network. Figure 2 presents the flow chart of cognitive lightweight-LR approach. Equations (1) and (2) present the logic explanation of linear logistic regression [54].

$$Y = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \cdots, \tag{1}$$

$$\operatorname{logit}(p) = \ln\left(\frac{p}{1-p}\right) = \beta 0 + \beta 1X1 + \beta 2X2 + \beta 3X3 + \cdots. \tag{2}$$

Figure 2 is the detailed flow chart regarding logistic regression. Initially, training data are used to formulate and train each function. Cost function is used to be calculated for logistic regression to test overall data. While, testing binary classification is utilized either "0" and "1" means "presence of attack" or "absence of attack" is identified easily.

5. Simulation Results

The simulation environment is designed for IoT-based UAV networks in anaconda python. UNSW-NB 15 dataset is used which consists of various cyberattacks such as DoS/DDoS, backdoors, fuzzers, exploits shellcode, and worms. The mentioned dataset consists of more than two million records. UNSW-NB 15 is a hybrid dataset where advanced data network traffic is incorporated. Three major problems can be easily tackled using UNSW-NB 15 dataset like low footprint, data traffic scenarios, and training/testing methods. However, for light weight algorithms the mentioned dataset are giving better results. Binary classification is utilized while simulating machine learning techniques which include decision tree, random forest, XGBoost, AdaBoost, bagging, and logistic regression [55–64]. Furthermore, the data are divided in training and testing modules which are as follows.

5.1. Data Training.

5.1. Data Training. Figure 3 provides detail information about training dataset. During training almost 56.06% data illustrates security attacks, while around 44.94% there is "no attack." Moreover, training dataset is quite balanced due to that false alarm is reduced.

5.2. Data Testing. Figure 4 shows data regarding testing dataset where 31.94% portion is for "no attack" scenario. However, 68.06% data are giving information regarding attacks.

Figure 5 depicts the detail information about training and testing datasets. The metric of high accuracy is maintained in optimal way using UNSW-NB 15 dataset. In high accuracy, there are two scenarios which include attack or no attack. Furthermore, if there will be attack but in reality no attack will be detected which will be false positive. Similarly, true negative will be having no attack where no attack can be identified.

The overall results of machine learning classifiers are presented in Figure 6. Logistic regression performs well in comparison with other algorithms. LR detects security attacks for about 82.54%, while, random forest 71.59%, XG Boost 49.54%, DT 49.17%, Bagging 44.70%, and AdaBoost around 28.39%. Also, Figure 6 provides information about the results of various machine learning classifiers in the area of IoT enabled FANETs. Figure 7 shows the similar results of Figure 6.

5.3. Comparative Discussion of ML-Based IDS. Table 2 elaborates the detailed comparison regarding ML-based intrusion detection system. The approach of network-based intrusion detection system is widely utilized. Also, anomaly-based IDS is quite popular approach to detect cyberattacks. In anomaly-IDS technique, a novel threshold is needed to be designed for identification of security attacks. While, signature-based IDS must have the concept of some possible attacks features stored in database. Although hybrid-IDS is the combination of anomaly and signature but the use is quite less. Therefore, the proposed solution is providing better possibilities to detect cyberattacks. In addition, Table 2 shows the studies which use to have information about different types of intrusion detection system. Also, machine learning-based IDS are widely utilized in the previous study.

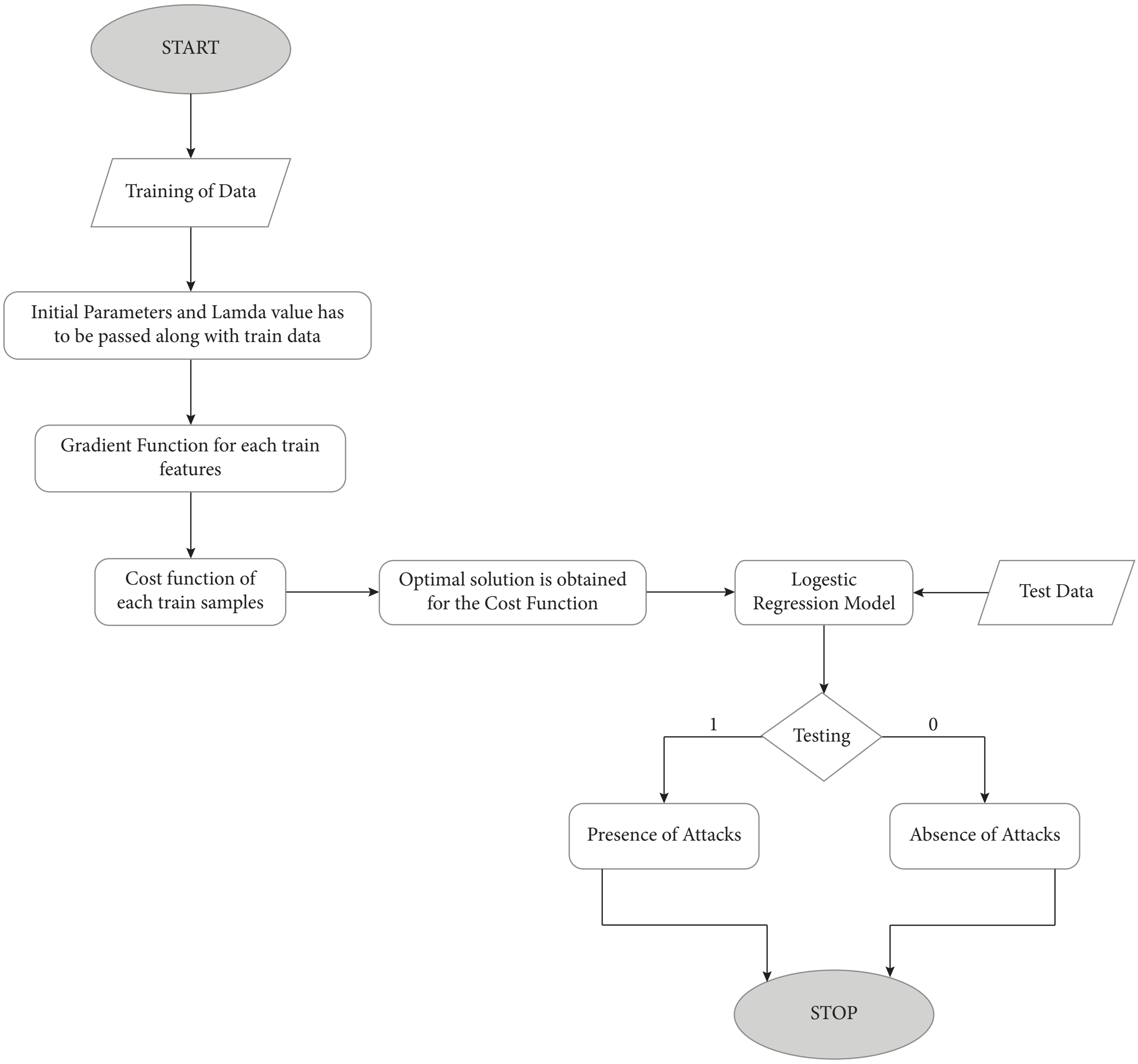
START
Training of Data
Initial Parameters and Lamda value has to be passed along with train data
Gradient Function for each train features
Cost function of each train samples
Optimal solution is obtained for the Cost Function
Logestic Regression Model
Test Data
Testing
1
0
Presence of Attacks
Absence of Attacks
STOP

Figure 2: Flowchart of cognitive lightweight-LR technique.

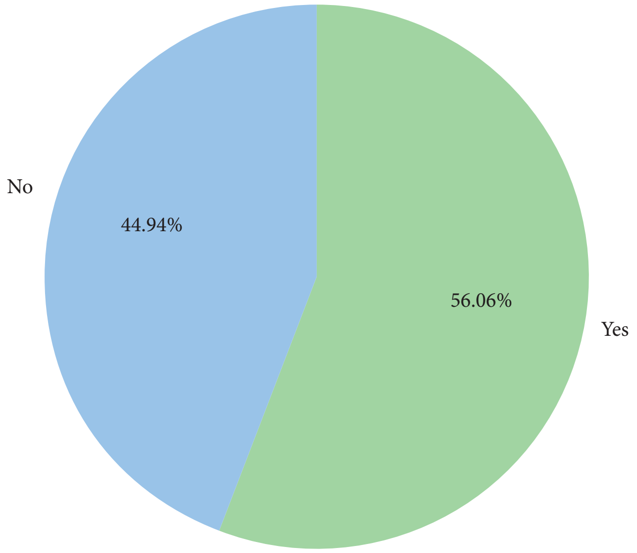

Figure 3: Training data for IoT-based UAV network.

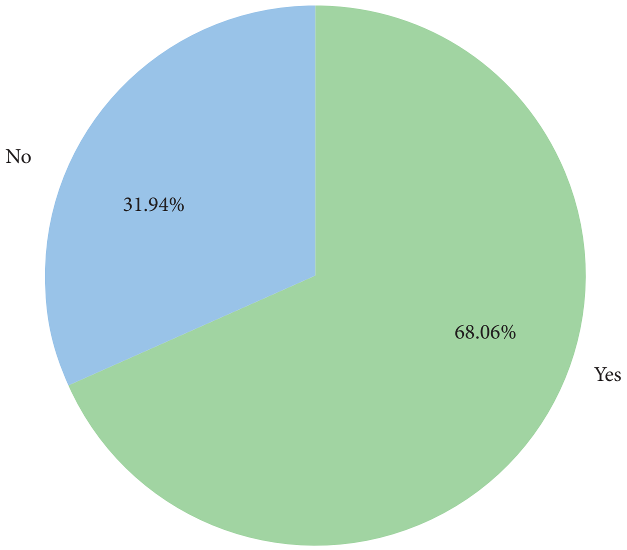

Figure 4: Testing data for IoT-based UAV network.

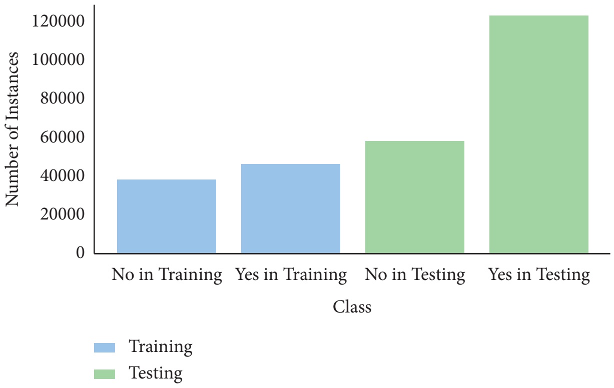

Figure 5: Comparison of training and testing datasets.

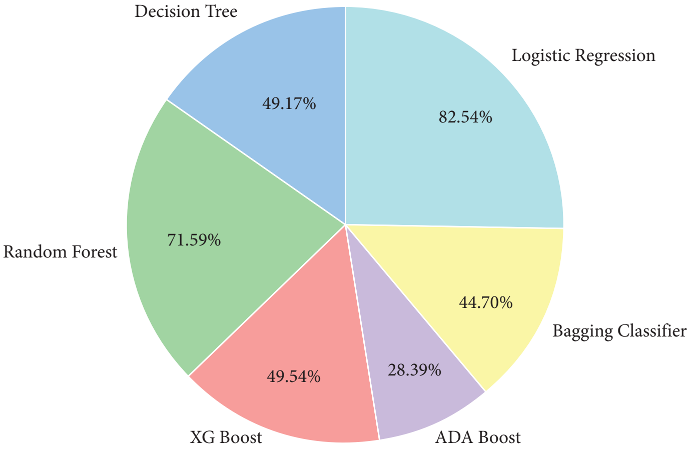

Figure 6: Performance analysis of machine learning classifiers (DT, RF, XGBoost, AdaBoost, Bagging and logistic regression).

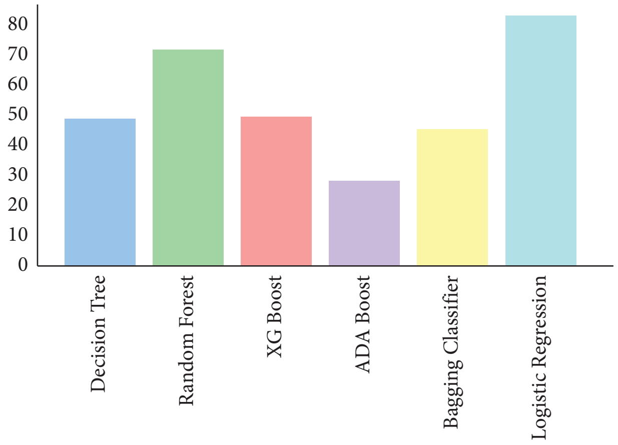

Figure 7: Comparative results of machine learning classifiers.

Table 2: Detailed comparative study of ML-based IDS.

Reference	Network-based IDS	Host-based IDS	Anomaly-based IDS	Signature-based IDS	Hybrid-based IDS	Machine learning-based IDS	Future scope
[65]	✓	X	✓	✓	X	✓	X
[66]	✓	X	✓	✓	X	✓	✓
[67]	✓	X	✓	✓	X	✓	✓
[68]	✓	X	✓	✓	X	✓	✓
[69]	✓	X	✓	✓	X	✓	X
[70]	✓	X	✓	✓	X	✓	✓
[71]	✓	X	✓	✓	X	✓	✓
[72]	✓	X	✓	✓	X	✓	X
[73]	✓	X	✓	✓	X	✓	✓
[74]	**X**	X	**X**	**X**	X	**X**	✓
[75]	✓	X	✓	✓	X	✓	X
[76]	✓	X	✓	✓	✓	✓	✓
Proposed work	✓	X	✓	✓	X	✓	✓

6. Conclusion

Machine learning-based techniques are deployed in IoT-based UAV networks. The main aim of this research study is to propose a novel concept of detecting abnormal behaviour using machine learning. Flying ad hoc networks is the combinations of group of UAVs formulate a network. FANET structure consists of UAVs, satellites, and ground-based stations. While, IoT sensor nodes are deployed on ground and UAVs use to collection information. However, cognitive lightweight-LR approach has reduced false alarm and balanced high accuracy in IoT-based UAV network. UNSW-NB 15 dataset is utilized to check the performance. Nowadays, security is one of the major concerns in almost every field of study. FANET-based IDS is the approach utilized to detect possible cyberattacks. The proposed approach has mimicked the overhead, and false data packets are detected easily. The simulation results shows that logistic regression performed better in comparison with other techniques. The concept of IoT-based UAV networks can be merged with smart cities in near future. In addition, optimization techniques and graph theory will give new directions to this study. Data traffic models and new datasets are the need of futuristic cities.

6.1. Future Direction

In near future, UAV network will be widely utilized for flying taxis in the concept of smart cities. Therefore, artificial intelligence, machine learning, deep learning, reinforcement-based learning, and federated learning can be utilized for intelligent IDS to detect cyberattacks. While in smart cities internet of everything will be used to advance communication. Routing protocols and communication standards need to be further investigated. Also, novel datasets need to be designed which will be helpful for researchers and scientists for further experimentations [77–79].

Data Availability

All the data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] A. Qayyum, L. Viennot, and A. Laouiti, "Multipoint relaying for flooding broadcast messages in mobile wireless networks," in *Proceedings of the 35th Annu. Hawaii Int. Conf. Syst. Sci*, pp. 3866–3875, Big Island, HI, USA, January 2002.

[2] I. U. Khan, I. M. Qureshi, M. A. Aziz, T. A. Cheema, and S. B. H. Shah, "Smart IoT control-based nature inspired energy efficient routing protocol for flying ad hoc network (FANET)," *IEEE Access*, vol. 8, pp. 56371–56378, 2020.

[3] I. U. Khan, A. Abdollahi, and A. Jamil, "Bisma baig, muhammad adnan aziz, and fazal subhan. "A novel design of FANET routing protocol aided 5G communication using IoT," *Journal of Mobile Multimedia*, vol. 27, pp. 1333–1354, 2022.

[4] J. Nzouonta, N. Rajgure, G. Wang, and C. Borcea, "VANET routing on city roads using real-time vehicular traffic information," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3609–3626, 2009.

[5] I. U. Khan, N. Z. Syeda Zillay, A. Abdollahi et al., "Reinforce based optimization in wireless communication technologies and routing techniques using internet of flying vehicles," in *Proceedings of the 4th International Conference on Future Networks and Distributed Systems (ICFNDS)*, pp. 1–6, St.Petersburg, Russian Federation, May 2020.

[6] J. Sun, F. Khan, J. Li, M. D. Alshehri, A. Ryan, and M. Wedyan, "Mutual authentication scheme for ensuring a secure device-to-server communication in the internet of medical things," *IEEE Internet of Things Journal*, vol. 2021, Article ID 3078702, 11 pages, 2021.

[7] I. U. Khan, A. Ryan, H. J. Alyamani et al., "RSSI-controlled long-range communication in secured IoT-enabled unmanned aerial vehicles," *Mobile Information Systems*, vol. 2021, Article ID 5523553, 11 pages, 2021.

[8] M. Ahmed and A. K. Pathan, "False data injection attack (FDIA): an overview and new metrics for fair evaluation of its countermeasure," *Complex Adaptive Systems Modeling*, vol. 8, no. 1, p. 4, 2020.

[9] I. U. Khan, A. Abdollahi, A. Ryan et al., "Intelligent detection system enabled attack probability using Markov chain in aerial networks," *Wireless Communications and Mobile Computing*, vol. 2021, Article ID 1542657, 9 pages, 2021.

[10] X. Sun, D. W. K. Ng, Z. Ding, Y. Xu, and Z. Zhong, "Physical layer security in UAV systems: challenges and opportunities," *IEEE Wireless Communications*, vol. 26, no. 5, pp. 40–47, 2019.

[11] H. Tran and C. So, "Enhanced intrusion detection system for an EH IoT architecture using a cooperative UAV relay and friendly UAV jammer," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 11, pp. 1786–1799, 2021.

[12] A. Abdollahi and M. Fathi, "An intrusion detection system on ping of death attacks in IoT networks," *Wireless Personal Communications*, vol. 112, no. 4, pp. 2057–2070, 2020.

[13] A. Pharate, H. Bhat, V. Shilimkar, and N. Mhetre, "Classification of intrusion detection system," *International Journal of Computer Application*, vol. 118, p. 7, 2015.

[14] R. A. Ramadan, A.-H. Emara, M. Al-Sarem, and M. Elhamahmy, "Internet of drones intrusion detection using deep learning," *Electronics*, vol. 10, no. 21, p. 2633, 2021.

[15] Q. Abu Al-Haija and A. Badawi, "High-performance intrusion detection system for networked UAVs via deep learning," *Neural Computing & Applications*, pp. 1–16, 2022.

[16] O. Bouhamed, O. Bouachir, M. Aloqaily, and I. Ridhawi, "Lightweight IDS for UAV networks: a periodic deep reinforcement learning-based approach," in *Proceedings of the 2021 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, pp. 1032–1037, IEEE, Bordeaux, France, May 2021.

[17] R. Shrestha, A. Omidkar, S. A. Roudi, R. Abbas, and S. Kim, "Machine-learning-enabled intrusion detection system for cellular connected UAV networks," *Electronics*, vol. 10, no. 13, p. 1549, 2021.

[18] A. Amouri, V. T. Alaparthy, and S. D. Morgera, "A machine learning based intrusion detection system for mobile Internet of Things," *Sensors*, vol. 20, no. 2, p. 461, 2020.

[19] M. A. Khan, M. A. Khan, S. Ullah Jan et al., "A deep learning-based intrusion detection system for MQTT enabled IoT," *Sensors*, vol. 21, pp. 7016–21, 2021.

[20] A. Ghaleb, F. Saeed, M. Al-Sarem et al., "Misbehavior-aware on-demand collaborative intrusion detection system using distributed ensemble learning for VANET," *Electronics*, vol. 9, no. 9, p. 1411, 2020.

[21] G. Carleo, I. Cirac, K. Cranmer et al., "Machine learning and the physical sciences," *Reviews of Modern Physics*, vol. 91, no. 4, Article ID 045002, 2019.

[22] A. E. Hassanien, A. Darwish, and S. Abdelghafar, "Machine learning in telemetry data mining of space mission: basics, challenging and future directions," *Artificial Intelligence Review*, vol. 53, no. 5, pp. 3201–3230, 2020.

[23] T. Leiner, D. Rueckert, A. Suinesiaputra et al., "Machine learning in cardiovascular magnetic resonance: basic concepts and applications," *Journal of Cardiovascular Magnetic Resonance*, vol. 21, no. 1, pp. 1–14, 2019.

[24] A. Mahmood and J.-L. Wang, "Machine learning for high performance organic solar cells: current scenario and future prospects," *Energy & Environmental Science*, vol. 14, no. 1, pp. 90–105, 2021.

[25] D. Soriano-Valdez, I. Pelaez-Ballestas, A. Manrique de Lara, and A. Gastelum-Strozzi, "The basics of data, big data, and machine learning in clinical practice," *Clinical Rheumatology*, vol. 40, no. 1, pp. 11–23, 2021.

[26] S. Lee, S. H. Lam, T. A. Hernandes Rocha et al., "Machine learning and precision medicine in emergency medicine: the basics," *Cureus*, vol. 13, p. 9, 2021.

[27] B. Charbuty and A. Abdulazeez, "Classification based on decision tree algorithm for machine learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, 2021.

[28] M. Pandey and V. K. Sharma, "A decision tree algorithm pertaining to the student performance analysis and prediction," *International Journal of Computer Application*, vol. 61, p. 13, 2013.

[29] B. Chandra and P. Paul Varghese, "Fuzzy SLIQ decision tree algorithm," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 38, no. 5, pp. 1294–1301, 2008.

[30] N. Bhargava, G. Sharma, R. Bhargava, and M. Mathuria, "Decision tree analysis on j48 algorithm for data mining," *Proceedings of international journal of advanced research in computer science and software engineering*, vol. 3, p. 6, 2013.

[31] M. Kumar, M. Hanumanthappa, and T. V. Suresh Kumar, "Intrusion Detection System using decision tree algorithm," in *Proceedings of the 2012 IEEE 14th international conference*

on communication technology, pp. 629–634, IEEE, Chengdu, China, November 2012.

[32] G. Biau and E. Scornet, "A random forest guided tour," *Test*, vol. 25, no. 2, pp. 197–227, 2016.

[33] M. Schonlau and R. Y. Zou, "The random forest algorithm for statistical learning," *STATA Journal*, vol. 20, no. 1, pp. 3–29, 2020.

[34] Y. Liu, Y. Wang, and J. Zhang, "New machine learning algorithm: random forest," in *International Conference on Information Computing and Applications*, pp. 246–252, Springer, Berlin, Germany, 2012.

[35] S. J. Rigatti, "Random forest," *Journal of Insurance Medicine*, vol. 47, no. 1, pp. 31–39, 2017.

[36] L. Zhu, D. Qiu, D. Ergu, Y. Cai, and K. Liu, "A study on predicting loan default based on the random forest algorithm," *Procedia Computer Science*, vol. 162, pp. 503–513, 2019.

[37] R. P. Sheridan, W. M. Wang, A. Liaw, J. Ma, M. Eric, and Gifford, "Extreme gradient boosting as a method for quantitative structure–activity relationships," *Journal of Chemical Information and Modeling*, vol. 56, no. 12, pp. 2353–2360, 2016.

[38] P. Carmona, F. Climent, and A. Momparler, "Predicting failure in the US banking sector: an extreme gradient boosting approach," *International Review of Economics & Finance*, vol. 61, pp. 304–323, 2019.

[39] Y.-C. Chang, K.-H. Chang, and G.-J. Wu, "Application of eXtreme gradient boosting trees in the construction of credit risk assessment models for financial institutions," *Applied Soft Computing*, vol. 73, pp. 914–920, 2018.

[40] R. Song, S. Chen, B. Deng, and L. Li, "eXtreme gradient boosting for identifying individual users across different digital devices," in *International Conference on Web-Age Information Management*, pp. 43–54, Springer, Berlin, Germany, 2016.

[41] T. Chengsheng, H. Liu, and B. Xu, "AdaBoost typical Algorithm and its application research," in *MATEC Web of Conferences*vol. 139, EDP Sciences, Article ID 00222, 2017.

[42] T.-K. An and M.-H. Kim, "A new diverse AdaBoost classifier," in *2010 International conference on artificial intelligence and computational intelligence*, vol. 1, pp. 359–363, IEEE, 2010.

[43] P. Wu and H. Zhao, "Some analysis and research of the AdaBoost algorithm," in *International Conference on Intelligent Computing and Information Science*, pp. 1–5, Springer, Berlin, Heidelberg, 2011.

[44] B. Sun, S. Chen, J. Wang, and H. Chen, "A robust multi-class AdaBoost algorithm for mislabeled noisy data," *Knowledge-Based Systems*, vol. 102, pp. 87–102, 2016.

[45] M. Zareapoor and P. Shamsolmoali, "Application of credit card fraud detection: based on bagging ensemble classifier," *Procedia Computer Science*, vol. 48, no. 2015, pp. 679–685, 2015.

[46] Sandag and A. Green, "A prediction model of company health using bagging classifier," *JITK (Jurnal Ilmu Pengetahuan Dan Teknologi Komputer)*, vol. 6, no. 1, pp. 41–46, 2020.

[47] E. Bauer and R. Kohavi, "An empirical comparison of voting classification algorithms: bagging, boosting, and variants," *Machine Learning*, vol. 36, no. 1, pp. 105–139, 1999.

[48] K. Machová, F. Barcak, and P. Bednár, "A bagging method using decision trees in the role of base classifiers," *Acta Polytechnica Hungarica*, vol. 3, no. 2, pp. 121–132, 2006.

[49] S. Kotsiantis and P. Pintelas, "Combining bagging and boosting," *International Journal of Computational Intelligence*, vol. 1, no. 4, pp. 324–333, 2004.

[50] D. Caigny, K. Arno, W. Koen, and D. Bock, "A new hybrid classification algorithm for customer churn prediction based on logistic regression and decision trees," *European Journal of Operational Research*, vol. 269, no. 2, pp. 760–772, 2018.

[51] A. Arabameri, B. Pradhan, K. Rezaei, M. Yamani, H. R. Pourghasemi, and L. Lombardo, "Spatial modelling of gully erosion using evidential belief function, logistic regression, and a new ensemble of evidential belief function–logistic regression algorithm," *Land Degradation & Development*, vol. 29, no. 11, pp. 4035–4049, 2018.

[52] D. Böhning, "Multinomial logistic regression algorithm," *Annals of the Institute of Statistical Mathematics*, vol. 44, no. 1, pp. 197–200, 1992.

[53] K. Chaudhuri and C. Monteleoni, "Privacy-preserving logistic regression," *Advances in Neural Information Processing Systems*, vol. 21, 2008.

[54] N. Srimaneekarn, H. Anthony, W. Liu, and C. Tantipoj, "Binary response analysis using logistic regression in dentistry," *International Journal of Dentistry*, vol. 2022, Article ID 5358602, 8 pages, 2022.

[55] N. Moustafa, G. Creech, and J. Slay, "Big data analytics for intrusion detection system: statistical decision-making using finite dirichlet mixture models," in *Data Analytics and Decision Support for Cybersecurity*, pp. 127–156, Springer, Cham, 2017.

[56] N. Moustafa, J. Slay, and G. Creech, "Novel geometric area analysis technique for anomaly detection using trapezoidal area estimation on large-scale networks," *IEEE Transactions on Big Data*, vol. 5, no. 4, pp. 481–494, 2019.

[57] N. Moustafa and S. Jill, "The evaluation of Network Anomaly Detection Systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set," *Information Security Journal: A Global Perspective*, vol. 25, no. 1-3, pp. 18–31, 2016.

[58] N. Moustafa and J. Slay, "UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set)," in *Proceedings of the 2015 Military Communications and Information Systems Conference (MilCIS)*, pp. 1–6, Canberra, ACT, Australia, November 2015.

[59] M. Sarhan, S. Layeghy, N. Moustafa, and M. Portmann, "Netflow datasets for machine learning-based network intrusion detection systems," 2020, https://arxiv.org/abs/2011.09144.

[60] N. Moustafa, B. Turnbull, and K. R. Choo, "An ensemble intrusion detection technique based on proposed statistical flow features for protecting network traffic of internet of things," *IEEE Internet of Things Journal*, vol. 6, no. 3, pp. 4815–4830, 2019.

[61] N. Moustafa, G. Misra, and J. Slay, "Generalized outlier Gaussian mixture technique based on automated association features for simulating and detecting web application attacks," *IEEE Transactions on Sustainable Computing*, vol. 6, no. 2, pp. 245–256, 2021.

[62] M. Keshk, N. Moustafa, E. Sitnikova, and G. Creech, "Privacy preservation intrusion detection technique for SCADA systems," in *Proceedings of the 2017 Military Communications and Information Systems Conference (MilCIS)*, pp. 1–6, Canberra, ACT, Australia, November 2017.

[63] N. Moustafa, G. Creech, and J. Slay, "Anomaly detection system using beta mixture models and outlier detection," in *Progress in Computing, Analytics and Networking*, pp. 125–135, Springer, Singapore, 2018.

[64] N. Moustafa and J. Slay, "A network forensic scheme using correntropy-variation for attack detection," in *IFIP*

International Conference on Digital Forensics, pp. 225–239, Springer, Cham, 2018.

[65] A. Ugendhar, S. Babu Illuri, R. Vulapula et al., "A novel intelligent-based intrusion detection system approach using deep multilayer classification," *Mathematical Problems in Engineering*, vol. 2022, Article ID 8030510, 10 pages, 2022.

[66] G. Thamilarasu and S. Chawla, "Towards deep-learning-driven intrusion detection for the internet of things," *Sensors*, vol. 19, no. 9, p. 1977, 2019.

[67] A. Dahou, M. A. Elaziz, S. A. Chelloug et al., "Intrusion detection system for IoT based on deep learning and modified reptile search algorithm," *Computational Intelligence and Neuroscience*, vol. 2022, Article ID 6473507, 15 pages, 2022.

[68] J. Ren, J. Guo, Q. Wang, Y. Huang, X. Hao, and J. Hu, "Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms," *Security and Communication Networks*, vol. 2019, Article ID 7130868, 11 pages, 2019.

[69] S. Einy, C. Oz, and D. N. Yahya, "The anomaly-and signature-based IDS for network security using hybrid inference systems," *Mathematical Problems in Engineering*, vol. 2021, Article ID 6639714, 10 pages, 2021.

[70] K. Albulayhi, Q. Abu Al-Haija, S. A. Alsuhibany, A. A. Jillepalli, M. Ashrafuzzaman, and F. T. Sheldon, "IoT intrusion detection using machine learning with a novel high performing feature selection method," *Applied Sciences*, vol. 12, no. 10, p. 5015, 2022.

[71] A. O. Alzahrani and J. F. A. Mohammed, "Designing a network intrusion detection system based on machine learning for software defined networks," *Future Internet*, vol. 13, no. 5, p. 111, 2021.

[72] P. Verma, D. Ankur, R. Singh et al., "A novel intrusion detection approach using machine learning ensemble for IoT environments," *Applied Sciences*, vol. 11, no. 21, Article ID 10268, 2021.

[73] H. Zainel and C. Koçak, "LAN intrusion detection using convolutional neural networks," *Applied Sciences*, vol. 12, no. 13, p. 6645, 2022.

[74] M. Althunayyan, N. Saxena, S. Li, and P. Gope, "Evaluation of black-box web application security scanners in detecting injection vulnerabilities," *Electronics*, vol. 11, no. 13, p. 2049, 2022.

[75] S. Ullah, J. Ahmad, M. A. Khan et al., "A new intrusion detection system for the internet of things via deep convolutional neural network and feature engineering," *Sensors*, vol. 22, no. 10, p. 3607, 2022.

[76] A. Khraisat, I. Gondal, P. Vamplew, J. Kamruzzaman, and A. Alazab, "Hybrid intrusion detection system based on the stacking ensemble of c5 decision tree classifier and one class support vector machine," *Electronics*, vol. 9, no. 1, p. 173, 2020.

[77] O. Friha, M. Amine Ferrag, L. Shu, L. Maglaras, K.-K. R. Choo, and M. Nafaa, "FELIDS: federated learning-based intrusion detection system for agricultural Internet of Things," *Journal of Parallel and Distributed Computing*, vol. 165, pp. 17–31, 2022.

[78] A. Duraisamy and M. Subramaniam, "Attack detection on IoT based smart cities using IDS based MANFIS classifier and secure data transmission using IRSA encryption," *Wireless Personal Communications*, vol. 119, no. 2, pp. 1913–1934, 2021.

[79] Loo, E. Chong, M. Y. Ng, C. Leckie, and M. Palaniswami, "Intrusion detection for routing attacks in sensor networks," *International Journal of Distributed Sensor Networks*, vol. 2, no. 4, pp. 313–332, 2006.