Detection of Sinkhole Attack in Wireless Sensor Networks Using Machine Learning

Samina Kousar[1], Humaira Ashraf[2], NZ Jhanjhi[3]

[1] International Islamic University Islamabad Pakistan; saminakousar491@gmail.com
[2] International Islamic University Islamabad Pakistan; humaira.ashraf@iiu.edu.pk
[3] School of Computer Science, SCS, Taylor's University, Subang Jaya, Malaysia, noorzaman.jhanjhi@taylors.edu.my

Corresponding Author:
Samina Kousar[1]
Email address: saminakousar491@gmail.com
Humaira Ashraf[2]
Email address: humaira.ashraf@iiu.edu.pk

Abstract

Wireless sensor networks are becoming more and more well-liked for establishing various communication systems. Throughout, there are many applications for wireless sensor networks. A wireless network's sensors are susceptible to a number of security vulnerabilities. Attacks through sinkholes are one of them, in which by providing false information about the routing path, it attracts nearby nodes. Instead of the base station, malicious nodes get data from sensor nodes. A malicious node serves as the base station in a sinkhole attack. This assault is cruel since it is hard to detect with in the network and resistant to many methods. This article provides a comprehensive summary of the literature on the topic of Wireless sensor networks for the detection of sinkhole attacks. To identify gaps in the literature, the current surveys are also examined. Existing techniques have a high false alarm rate which leads to low accuracy and high energy consumption for sinkhole attack detection in a wireless sensor network. Numerous current methods utilizing a variety of methodologies are also critically evaluated in terms of delay, detection rate, packet delivery ratio, and energy consumption. Additionally, our study examined the unsolved problems in the field of identification of sinkhole attacks in the wireless network of sensors. We propose sinkhole attack detection with machine learning (SAD_ML) technique with less energy consumption and high detection accuracy for the classification of sinkhole attack in the wireless sensor network. For simulation of the proposed SAD_ ML method for sinkhole attack detection in the wireless sensor network use MATLAB simulator. First, we find suspicious nodes by using AODV (Ad hoc on Demand Distance Vector) protocol and ACK method. Secondly, suspicious nodes classify by using machine learning classification algorithms for sinkhole attack detection in the wireless sensor network. In terms of detection accuracy, the comparison of machine

learning classification algorithms reveals that SVM results are better than other ML models. Our SAD_ ML technique accuracy is 96 percent with the SVM algorithm.

Keywords: sinkhole attack; wireless sensor network; network security attacks; AODV; attack detection

Introduction

The A Wireless Sensor Network (WSN) is a system of device nodes that gathers data about an area's environment. Environmental factors like temperature, humidity, sound, wind, motion, pressure, pollution, vibration, and more are measured via wireless sensor networks [3]. The user interacts with the network via a base station or sink. The sensor nodes interact with one another through radio waves. The wireless sensor nodes collect data about the environment around them and respond to commands from the control site by carrying out specific tasks or sending sensing samples. The nodes' processing power, communication bandwidth, and storage capacity are all limited. WSNs have become a defining technology for the future because of the wide range of industries it may be used in encompassing law enforcement, medicine, industrial management and supervision, and numerous more fields [8]. Wireless networks are susceptible to security attacks because of the features of the wireless broadcast medium. WSNs are also susceptible because of the resource limitations in the sensor nodes. The sensor nodes can also be installed in an unattended setting, which leaves them physically unprotected and open to enemy capture. As a result, WSN security emerges as a crucial problem that engages lots of scholars. Numerous attacks, including sinkhole attacks, HELLO flood attacks, Denial of Service attacks, blackhole attack, selective forwarding attacks, Sybil attacks, and wormhole attack take advantage of weaknesses in WSNs [21]. Figure 1 shows a sinkhole attack, in which by providing false information about the routing path, it attracts nearby nodes. Data is sent by nodes to the malicious node rather than the base station, malicious nodes get data from sensor nodes. During a sinkhole attack, in which a malicious node acts as the base station. This assault is cruel since it is hard to detect within the network and resistant to many methods. The effectiveness of the network may be decreased by a sinkhole attack by misleading nearby nodes and enabling them to launch additional attacks. A number of packets sensed by sinkhole attacks may be dropped or altered as they flow through the affected node. Although numerous approaches to detecting sinkhole attacks in WSNs have been put forth in the literature, high energy consumption and low accuracy of the detection of the sinkhole attacker remain serious issues. We propose sinkhole attack detection with machine learning (SAD_ML) technique with less energy consumption and high detection accuracy identification of sinkhole attack in wireless sensor network. First, we find suspicious nodes by using the AODV (Ad hoc on Demand Distance Vector) protocol and ACK method. Secondly, suspicious nodes classify by using machine learning classification algorithms for identification of sinkhole attack in the wireless sensor network.

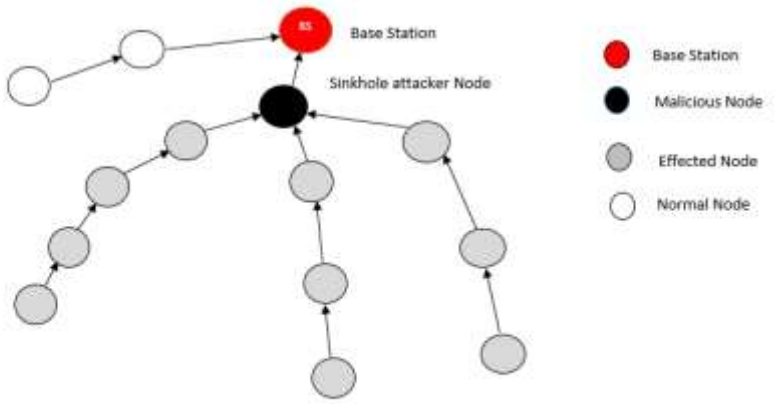

Figure 2. Sinkhole Attack in Wireless Sensor Network

Aim And Objective

Aim: Detection of Sinkhole attack detection in wireless sensor network

Objective: Detection with a low false alarm rate leads to the high accuracy of Sinkhole attack detection in wireless sensor network . Low energy consumption for Sinkhole attack detection in wireless sensor network

Primary Contribution

The following are the primary contributions of this article:

- A thorough analysis of the issues with existing methods was conducted for identification of sinkhole attacks.
- In this article, sinkhole attack detection using the machine learning (SAD_ML) technique in wireless sensor networks is proposed with improved accuracy and reduced energy consumption.

Paper Structure

Paper Structure Section 2 presents a literature review of related identification of sinkhole attack techniques in WSN schemes. The 3 section gives a suggested methodology for this research paper. Section 4, result analysis of the proposed method identification of Sinkhole attack in the wireless sensor network. Section 5 discusses the conclusion and future work.

Literature Review

In this section, several sinkhole attack detection techniques are discussed. The trust-based, lightweight RF Trust model presented by Prathapchandran & T [1]offers a method to guarantee the safety of an internet of things.It was intended to identification of the sinkhole attack in IoT scenarios that use RPL (Routing Protocol for Low power and Lossy networks) protocol. Identifying and eliminating sinkhole nodes from the network, improves trustworthy path in the internet of things context. Suggested method employs Random Forest (RF) to prevent sinkhole attacks and boost network performance. The suggested work's merits have a high packet delivery ratio (PDR), high throughput, and minimal average delay. Low precision in detecting sinkhole attacks is the drawback.

Yadav & Tak [2] a new method that is based on the analysis of routing behavior was proposed as a methodology for a wireless sensor network (WSN) to detect sinkhole attacks. The AODV reactive routing technology eliminates the need to include the source route in each packet. This means that it is simpler and more effective than other routing protocols during packet transfer. Al Maslamani and Mohamed [3] proposed and developed a detection approach against sinkhole attacks with the SI (Swarm Intelligence optimization) algorithm. The proposed method combines the ABC (Artificial Bee Colony) optimization algorithm and weight estimation mechanism to enhance the accuracy results of sinkhole assaults. To evaluate the effectiveness of the suggested work on the basis of convergence speed, energy consumption, packet overhead, detection time, and detection accuracy, comprehensive simulations have been done. The results show that the proposed mechanism can identify sinkhole attacks with a high percentage of classification accuracies.

Jamil et al.[4] suggested a method in order to identify and prevent sinkhole attacks in the network, a new Secured routing protocol for low energy called the CLS-RPL (Cross Layers Secured RPL). The RPL routing protocol has been improved by this routing protocol. A cross-layer routing protocol called CLS-RPL utilizes data from the data connection layer as part of its security mechanism. In order to detect sinkhole attacks, CLS-RPL employs a novel method and concept called overhearing, which enables a child node to listen in on its parent transmission. A straightforward security method used by CLS-RPL offers a high packet delivery ratio. The results demonstrate that when compared to the RPL protocol, CLS-RPL offers a 52% increase in terms of packet delivery ratio. Nithiyanandam and Latha [5] have a reliable approach to detecting sinkholes in wireless acoustic sensor networks based on voting method. In order to effectively generate a Boolean key decreased to distribute suspect list among nodes that have been alerted, they employed the Artificial Bee Colony technique. Good results were obtained from the proposed algorithms in PDR (Packet Delivery Ratio). Mehta and Jasminder [6] provided two novel techniques: Removal of Highest Severity Node (RHSN) And Severity Detection of Sinkhole Attack (SDSN) for removing the malicious node for identification. According to the results, the suggested technique outperformed the alternatives on the basis of delay, throughput, packet loss, and energy use. This technique has high energy utilization.

Jatti and V [7] proposed an agent-based method for sinkhole attack detection and prevention. In this approach, agents are utilized to negotiate three times and give each node information from its reliable peers. The merits of the method were high packet delivery ratio and high throughput. In wireless sensor networks, Nithiyanandam and P [8] proposed using artificial bee colonies to detect sinkholes. This technique compares the node IDs defined in the rule set to identify the compromised node. By reducing the total amount of time required to identify the affected node, artificial bee colonies the decrease lower the packet loss and raise PDR (packet delivery ratio) percentile. Nwankwo and Shafi [9] offer an ant colony optimization-enhanced sinkhole detection technique. To enhance identification of sinkhole attack via packet loss, PDR(packet delivery rate), energy exchange, and throughput in a wireless sensor network.Tak and Ashish [10] suggested using a secure AODV protocol to fid sinkhole attacks in wireless sensor networks. In this work, five factors, delay, PDR, Normalized Routing Load, Routing Overhead, and Average throughput —were taken into account.

Mondal et al. [11] a detection method for sinkhole attacks was suggested that makes use of the formula Euclidean distance between base station and each node. Main benefit of proposed suggested method was not need any additional communication expenses or hardware setup. New Mutual Authentication Scheme by Kumar and Nitika [12] Isolates Sinkhole Attack in the network of wireless sensors . Find out network per-hop delay using this method. Isolate the node that is causing the delay if it reaches 2 ms. The

outcomes showed that, in comparison to other strategies, the newly proposed algorithm functioned more successfully. By using a forward chaining inference engine, An and Tae [13] propose a method to identify sinkhole assaults and set limits and variations as principles of the level of knowledge.They do this using a specification-based approach to intrusion detection. The suggested technique improved the sinkhole detection accuracy percentage by an average of seven percentage when compared to the current method.

In wireless sensor networks, Raj and Darpan [14] introduced Sink Hole Attack identification using Two-Step Verification Technique. It was initially verified that there were any malicious nodes in the network. The second stage was focused on finding the network's malicious node. To analyse the findings, many aspects including latency, throughput, and packet loss, were taken into account. This showed that the suggested technique performs better than the present scheme to find malicious nodes in the network. Energy-saving method to find Sinkhole assaults Utilizing Roving identification Internet Protocol version 6 over low-power wireless networks was proposed by Pradeep Kumar et al. [15]. This paper's main objective was to address routing attacks. They would like to identify sinkhole attacks in the Internet of Things ecosystem wtheirthey designed energy-efficient, lightweight protection solutions. The findings show that the technique was straightforward., performs admirably in identifying sinkhole attacks, and offers rates of 85.66% Tand PR, and 84.21% TNR. Additionally, it displays the minimal RAM/ROM utilization and energy usage, which are comparable.

An intrusion detection approach was suggested by Ahmed et al. [16] to defend the IoT infrastructure from sinkhole assaults. A paradigm that uses richly equipped and trained edge nodes to exchange messages to find different types of sinkhole attacker nodes. The model was put into use in practice using a well-known NS2 simulator. The suggested approach had a detection results of more than 85% through false-positive rate only 1.4%, which was superior to the schemes previously suggested. The proposed plan was also a good fit for a critical platform like a security and monitoring system. The suggested method of detecting Blockhole, Wormhole, and Cooperative black hole threats in IoT networks cannot be put into practice. For the network of agriculture-basics WSN, Iman et al. [17] suggested a sinkhole assault detection system for intrusions. By integrating in Internet of Things sensors into various environments. Many farmers used WSN to streamline the process of monitoring and gathering crucial data about the state of their farms and greenhouse to maintain the appropriate level of humidity, temperature, and light exposure. This suggested paper has presented an intrusion detection system (IDS). Three network topology simulations were used in the project to compare the results depending depending on how well the network traffic performs. The results demonstrated that, in comparison to sinkhole networks without IDS and networks free from attacks, the agriculture WSN will perform better when IDS was included. Therefore, it shows that the suggested IDS could identify the network when unusual behavior showed up in the network topologies.

Depending on the Neighbor Density Estimation Technique put out by Karthigadevi et al. [18], increased wireless sensor networks' ability to detect and stop sinkhole assaults. To find and stop the sinkhole threat, a unique decentralized use the network density estimate technique to find sinkholes mechanism was proposed. In order to keep track of information about its neighbours, every node inside a network keeps a neighbor table, and each node uses this neighbor table to perform the Network Density Estimation Technique. Each node in a network records information about its neighbors. By utilizing all of these techniques for network density estimation, one may determine the network density and determine whether any malicious nodes were present in the area. The adjacent nodes were informed of the identified malicious nodes so that they can ignore them during subsequent broadcasts. The overhead

of gathering snapshots and routes was decreased by this technique. By boosting the volume of Best Effort traffic, this technique boosts network throughput.

The strategy suggested by Urvashi et al. [19] uses node-localization, with the base station examining hop latency. A specific quantity of sensor nodes were deployed over clusters the network in a specific size. Algorithms for location-based clustering partition the network into groups. . The base station will identify and isolate malicious nodes using the node localization approach. The base station gathers data about the node's position using the node localization method. The data also includes each of their separations from the base station. A network node that can increase the maximum delay times will be recognized as malicious. To identify sinkhole attacks in a network of wireless sensors, Kenneth et al. [20] present an extended ant colony optimization approach (EACO). Small sensor nodes that can perceive and interpret information made of Networks Wireless Sensor. When an adversary wireless sensor network node impersonates the real node closest to the ground station, all data passes through, giving the attacker the chance to alter, delete, or postpone information being sent to the sink node. This procedure describe, they provide an ant colony optimization-enhanced sinkhole identification technique to enhance sinkhole detection throughput, packet delivery rate and packet drop, and exchange of energy in wireless sensing networks. WSN security emerges as a crucial problem that engages lots of scholars [21]. Numerous attacks, including sinkhole attacks, HELLO flood attacks, Denial of Service attacks, blackhole attack, selective forwarding attacks, Sybil attacks, and wormhole attack take advantage of weaknesses in WSNs. In our systematic study of a secure blockchain-based decentralized scheme for the Internet of Things (IoT), we draw upon the foundational knowledge established in [24-38].

Materials and Methods

We propose sinkhole attack detection in the wireless sensor network with machine learning (SAD_ML) technique with less energy consumption and high accuracy scheme for the classification of sinkhole attack detection in the wireless sensor network. In Figure2, First, we find suspicious nodes by using the AODV (Ad hoc on Demand Distance Vector) protocol and ACK method. Secondly, suspicious nodes classify by using machine learning classification algorithms for sinkhole attack detection in the wireless sensor network.

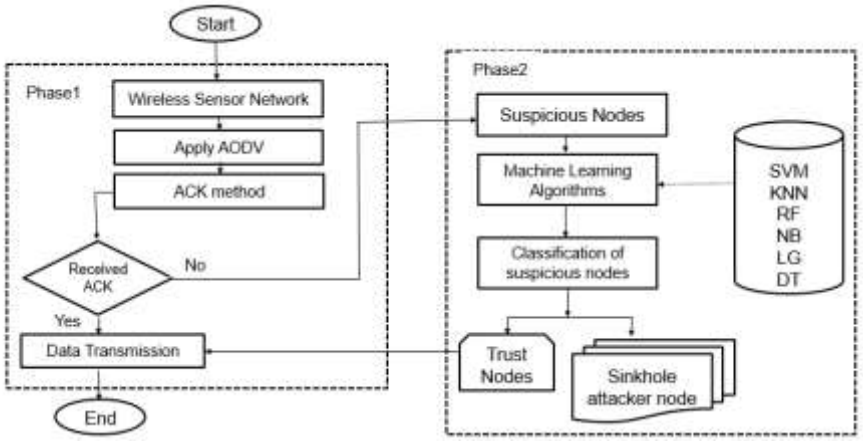

Figure 2. proposed methodology for sinkhole attack detection in WSN

Table 1: Notation for Scheme Algorithm

Symbols	Description
BS	Base station
SSNs	Suspicions Nodes
PDR	Packet Delivery Ratio
EC	Energy Consumption
RREQ	Rout Request
RREP	Rout Reply Packet
TTL	Time to Live
SHN	Sinkhole Node
SSNs	Suspicions Nodes
ACK	Acknowledgment

Phase 1(Find suspicious nodes) In this phase we find suspicious node by using the AODV (Ad hoc on Demand Distance Vector) protocol and ACK method through step by step in algorithm 1. Step 1: In wireless Sensor Network each node send data to sinkhole node. Step 2: Sensor node sends Route Request (RREQ) to the sinkhole node with Time To Live (TTL), hop count, node ID . Step 3: When the Sinkhole node received Route Request (RREQ) of packet then store data of source node. Step 4: After received Route Request (RREQ) Sinkhole node send Route Reply (RREP) to the sinkhole node with Time To Live (TTL), hop count, node ID . Step 5: If source node received acknowledgment ACK from sinkhole node within Time To Live (TTL) and hop count. Then sinkhole node is trust normal node. Step 6: If source node not received acknowledgment ACK from sinkhole node within Time To Live (TTL) and hop count. Then sinkhole node is sinkhole attacker.

Alorithm 1 Find suspicious node in wireless sensor network

1. **Begin**
2. **For** each SNs € SHN **do**
3. All SNs send data toward with RREQ to the sinkhole node with TTL, hop count, ID in a packet of each sensor nodes
4. The SHN collects and stores data received from SNs
5. Then SHN send ACK in term of RREP to SNs within TTL and hop count
6. SNs check ACK and hop count of SHN is reliable in term RREP
7. **If(**SN received ACK from SHN within TTL and hop count) **then**
8. SNs € SHN
9. **Else if (**SN received no ACK from SHN within TTL and hop count) **then**
10. SNs € SSNs
11. **End if**

12. **End for**
13. **End**

Phase 2 (Detection of maliicious node in wireless sensor network) Suspicious nodes classify by using machine learning classification algorithms for sinkhole attack detection in the wireless sensor network in algorithm 2 describe step bt step. Step1: access suspicious nodes Step2: Classification of trusted nodes and sinkhole attacker nodes from suspicious nodes Step3:Check threshold of Packet Delivery Ratio (PDR), Delay, Energy Consumption (EC), Honesty, Distance is hop count from base station. Step4: If Packet Delivery Ratio(PDR) is greater than threshold then Packet Delivery Ratio(PDR) is Good. Step5: If Packet Delivery Ratio(PDR) is less than threshold then Packet Delivery Ratio(PDR) is Bad. Step6: If Delay is greater than threshold then Delay is High. Step7: If Delay is less than threshold then Delay is low. Step8: If Energy Consumption (EC) is greater than threshold then Energy Consumption(EC) is Heavy. Step9: If Energy Consumption(EC) is less than threshold then Energy Consumption(EC) is Normal. Step10: Honesty is packet transmission is successful or not wireless sensor network. If Honesty is greater than threshold then Honesty is Yes. If Honesty less than threshold then Honesty is No. Step11: Distance is hop count from sinkhole node. If distance is greater than threshold then Distance is More. If Distance less than threshold then Distance is Less. Step12: Which nodes have good Packet Delivery Ratio(PDR), low delay, normal Energy Consumption(EC), honest has yes and less Distance from Base Station are trusted Normal nodes. Step13: Which nodes have bad Packet Delivery Ratio (PDR), high delay, heavy Energy Consumption(EC), honest has no and more Distance from Base Station are sinkhole attacker.

Alorithm 2 Detection of maliicious node in wireless sensor network

1. **Begin**
2. **For** access suspicious nodes **do**
3. Classification of trusted nodes and sinkhole attacker nodes
4. Check threshold of PDR, Delay, EC, Honesty, Distance from base station
5. **if(**PDR>threshold) **then**
6. PDR=Good
7. **else if**(PDR < threshold)**then**
8. PDR=Bad
9. **end if**
10. **if(**Delay > threshold) **then**
11. Delay = High
12. **else if**(Delay < threshold)**then**
13. Delay = Low
14. **end if**
15. **if(**EC > threshold) **then**
16. EC = Heavy
17. **else if**(EC < threshold)**then**
18. EC = Normal

19. **end if**
20. **if(**Honesty > threshold) **then**
21. Honesty = Yes
22. **else if**(Honesty < threshold)**then**
23. Honesty = No
24. **end if**
25. **if(**Distance > threshold) **then**
26. Distance = More
27. **else if**(Distance < threshold)**then**
28. Distance = Less
29. **end if**
30. **if**(PDR=Good, Delay=low, EC=normal, Honest=yes, Distance=less)
31. **then**
32. Trust Normal Nodes
33. **Else if(**PDR=Bad, Delay=high, EC=heavy, Honest=no, Distance=more)
34. Sinkhole Attacker Nodes
35. **end if**
36. **end for**
37. **end**

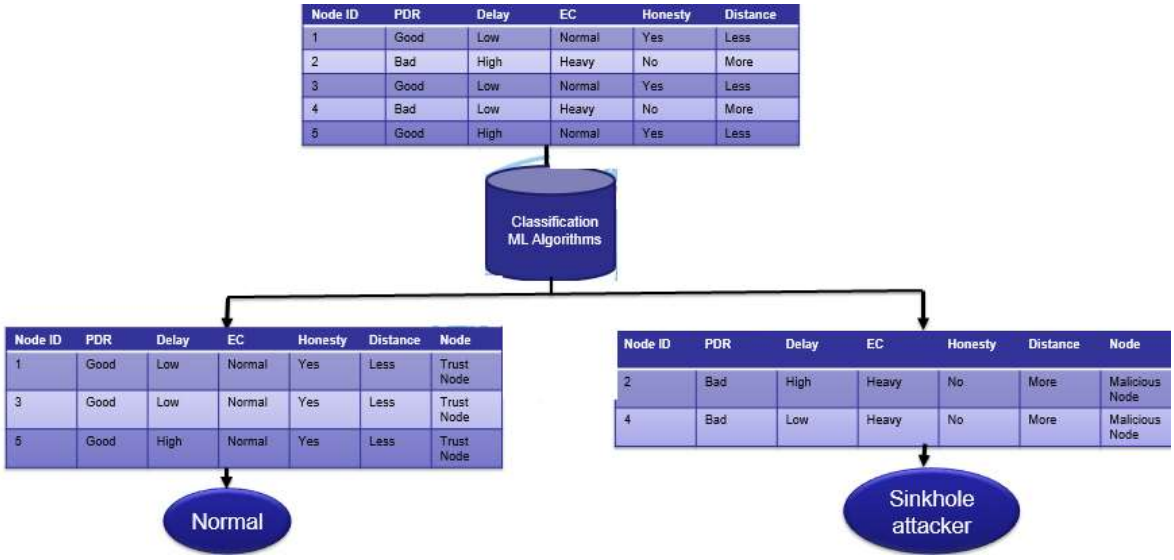

Figure 3. Factors used for sinkhole attack detection in WSN

In figure 4 shows the implementation of machine leaning models are Random forest, SVM, Decision Tree, Naïve Base, KNN, and Logistic Regression, and KNN for classification of sinkhole attacker.

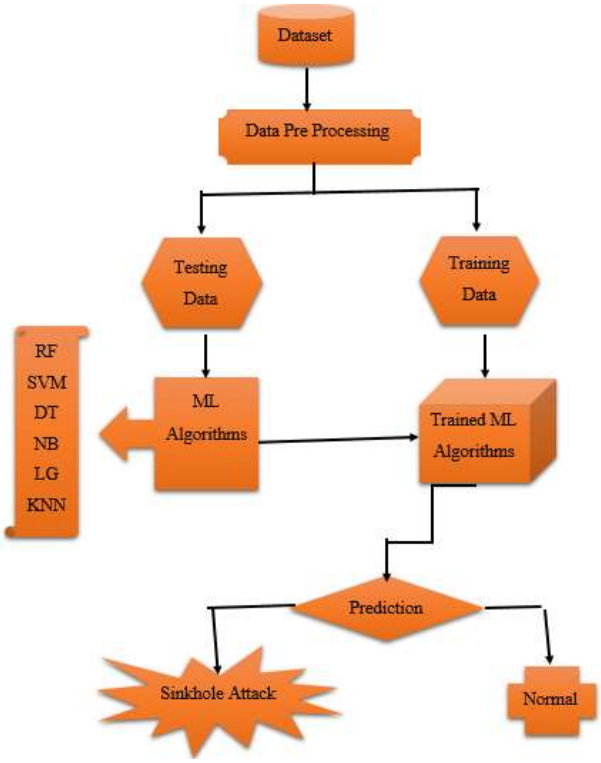

Figure 4. Machine Learning Model Implementation

Mathematical Modeling for SAD_ML Technique

Packet transmission between nodes of the wireless sensor network. In which sensor node forward route requests (RREQ) to the destination node in the wireless sensor network. The destination node receiving an RREQ then replies with a route reply (RREP) packet that is routed back to the original source.

SRREQ = Hop-count s +Time to live s+NIDs (1)

DRREP = Hop-count d+time to lived +NIDd (2)

Where,

$\mathbf{S_{RREQ}}$ = Source node send rout request to designation node

$\mathbf{D_{RREP}}$ = Destination node send rout reply to source node

TTL = Time-to-live (TTL) is a value for the period of time that a packet network

NID = Node identification

A packet of source node i forward with Time To Live(TTL) and Hop account toward destination node j. After receiving the packet destination node j reply to source node i within Time To Live(TTL) and Hop account.

Packet acknowledgment is received from destination node j within Time To Live (TTL) and Hop account. When a packet of source node i forward with Time To Live(TTL) and Hop account toward destination node j. After receiving the packet destination node j replies Packet acknowledgment to source node i within Time To Live(TTL) and Hop account. If the acknowledgment is not received within time

to live and Hop count then it is a suspicious node. If the acknowledgment is received within time to live and Hop count then it is not suspicious node.

ACKrec= **DRREP(Nj€i)yes** **(3)**

Where,

ACK_{rec} = acknowledgment received

D_{RREP} **=** Destination node send rout reply to source node

When Packet send RREQ to destination node then destination node Reply rout RREP to source node. If source node not received ACK from destination node with time to live and hop count then node is suspicious node.

Nsus=NO ACKrec**(j€i)** **(4)**

Where,

N_{sus} **=** suspicious node

ACK_{rec} = acknowledgment received

ACK_{no} = No acknowledgment received

The packet delivery ratio is good when it is greater than the threshold. The packet delivery ratio is bad when it is smaller than the threshold.

PDR_{good} = PDR>threshold **(5)**

PDR_{Bad} = PDR < threshold **(6)**

Where,

PDR_{good} = Good packet delivery ratio

PDR_{Bad} = Bad packet delivery ratio

Delay is high when the predefined threshold value is greater. When the delay is greater than the predefined value threshold then the delay is low.

$$Delay_{High} = \text{Delay} > \text{threshold} \quad \textbf{(7)}$$

$$Delay_{Low} = \text{Delay} < \text{threshold} \quad \textbf{(8)}$$

Where,

$Delay_{High}$ = High Delay

$Delay_{Low}$ = Low Delay

Energy Consumption (EC) is Heavy when Energy Consumption (EC) is greater than threshold value. Energy Consumption (EC) is normal when Energy Consumption (EC) is smaller than threshold value.

$$EC_{Heavy} = EC > \text{threshold} \quad \textbf{(9)}$$

$$EC_{Normal} = EC < \text{threshold} \quad \textbf{(10)}$$

Where,

EC_{Heavy} = Energy Consumption (EC) is Heavy than threshold value.

EC_{Normal} = Energy Consumption (EC) is normal) is smaller than threshold value.

Honesty of packet transmission is successful or not wireless sensor network. Honesty is yes when packet transmission is a successful wireless sensor network. Honesty is yes when it is greater than a threshold value. Honesty is no when it is smaller than a threshold value.

$$Honesty_{(Yes)} = \text{Honesty} > \text{threshold} \quad \textbf{(11)}$$

$$Honesty_{(No)} = \text{Honesty} < \text{threshold} \quad \textbf{(12)}$$

Where,

$Honesty_{(Yes)}$ = Honesty is yes when it is greater than a threshold value. $Honesty_{(No)}$= Honesty is no smaller than a threshold value

Distance is hop count from the source node to the destination node. Distance is more when distance is greater than a threshold value. Distance is less when distance is smaller than a threshold value.

$$Distance_{more} = \text{Distance} > \text{threshold} \quad \textbf{(13)}$$

$$Distance_{less} = \text{Distance} < \text{threshold} \quad \textbf{(14)}$$

Where,

$Distance_{more}$ = Distance is more when distance is greater than a threshold value.
$Distance_{less}$ = Distance is less when distance is smaller than a threshold value

Trust node is when PDR is Good, Delay is low, EC is normal, Honest is yes and Distance is less. Sinkhole attacker node is when PDR is Bad, Delay is high, EC is heavy, Honest is no and distance is more.

TN = PDR=Good, Delay=low, EC=normal, Honest=yes, Distance=less **(15)**

SHN_{attack} = PDR=Bad, Delay=high, EC=heavy, Honest=no, Distance=more **(16)**

Where,

TN = Trusted Node

SHN_{attack} = Sinkhole attacker node

Implementation of SAD_ML Scheme

This result analysis section describes Sinkhole attack detection in Wireless Sensor Networks using the machine learning SAD_ML technique of results obtained by data testing, training, and simulation. The simulation performance measurement of the suggested SAD_ML model has compared the Random Forest Trust approach [1] in terms of energy consumption, and accuracy.

Dataset Creation

Jupyter Notebook tool is used for the creation of the dataset for the simulation of proposed SAD_ML modal Sinkhole attack detection in Wireless Sensor Networks using machine learning. The selected attributes of the dataset are described in Table 2.

Table 2. dataset Attributes

Variables	Values
Number of Feature	08
Categories	2(Normal,Sinkhole_Attacker)
No. of samples	4000
Normal data	5000
Malicious data	800
Training size	75%
Testing size	25%

Figure 5 shows the generation of dataset for sinkhole attach detection in wireless sensor networks.

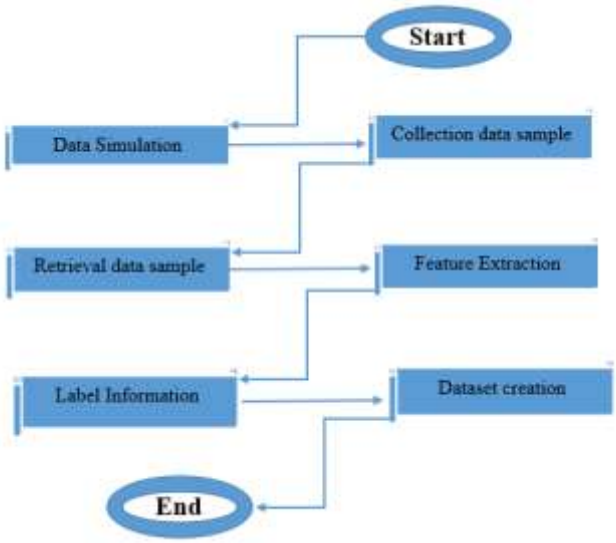

Figure 5. Dataset Generation

Simulation Setup

The proposed SAD_ML is evaluated by performing in MATLAB simulator. The parameters of the simulation used in the evaluation are presented in Table 3. The wireless sensor network's secure routing protocol is used AODV protocol.

Table 3 Simulation Parameter of proposed ML-SAD Modal

Parameters	Description
Network Type	Wireless
Network area	1000 x 1000 m
Sinkhole Node location	890 x 690 m
Number of nodes	100
Initial energy	1J
Packet size	78 bytes
Routing protocol	AODV
Simulation time	3600 s

Simulation Results

Suspicious nodes detection

Figure 7 shows the to identify suspicious nodes we use the Ad hoc on Demand Distance vector (AODV) protocol and ACK-based method in Wireless Sensor Network.

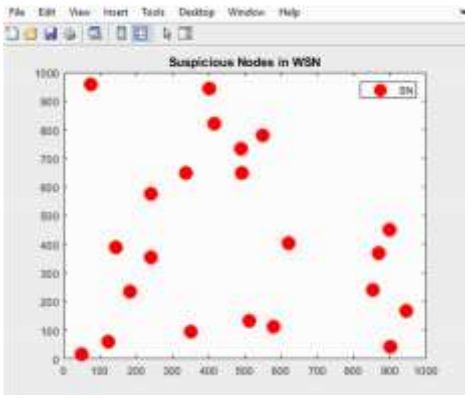

Figure 7. Suspicious nodes

One of the protocols developed to overcome several significant performance-related problems is AODV. It is a protocol for improvement. To the target node in the wireless sensor network, it transmits route requests (RREQ). After receiving an RREQ, the destination node responds with a route reply (RREP) packet that is sent back to the source. These nodes are suspicious nodes since they have not received acknowledgement.

In figure 8 shows the classification of trusted nodes and sinkhole attacker nodes from suspicious nodes. We use feature Check threshold of Packet Delivery Ratio (PDR), Delay, Energy Consumption (EC), Honesty, Distance is hop count from base station. Which nodes have bad Packet Delivery Ratio (PDR), high delay, heavy Energy Consumption (EC), honest has no and more Distance from Base Station are malicious nodes.

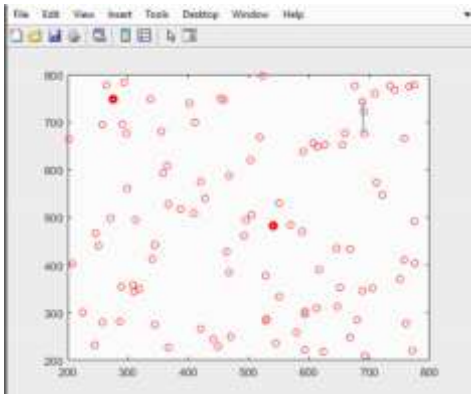

Figure 8. Sinkhole attacker nodes

Accuracy of Sinkhole attack detection in Wireless Sensor Networks using the machine learning shown in figure 9 with different machine learning algorithms.

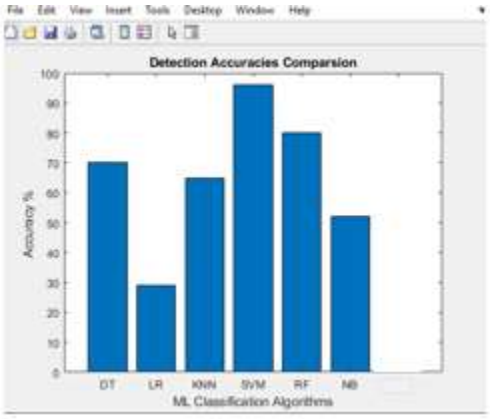

Figure 9. Detection Accuracy comparison

Figure 10 shows the energy consumption of nodes in the wireless sensor network. Nodes consume energy in joule

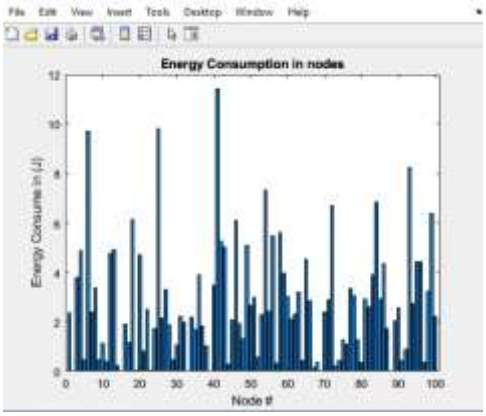

Figure 10. Energy consumption in nodes

Figure 11 shows the Time to live of packets of nodes in the wireless sensor network. Nodes of time of live in seconds.

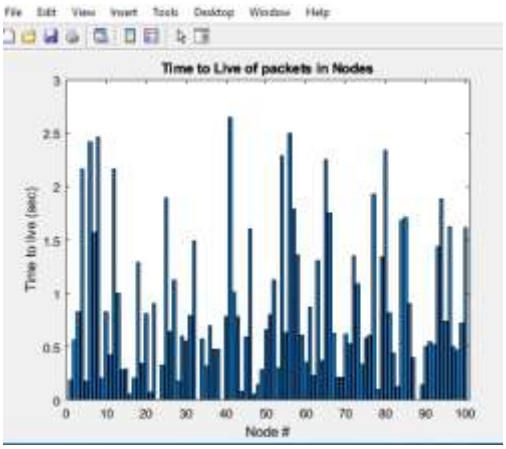

Figure 11. Time to live of packets in nodes

Comparison of results with literature

Figure 12 shows the detection of sinkhole attack in a wireless sensor network with SAD_ ML technique accuracies comparison of results with the literature.

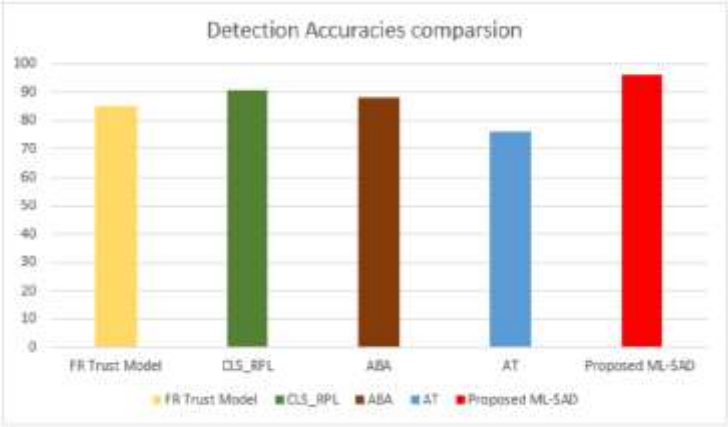

Figure 12. Detection accuracies comparison with literature

Figure 13 shows the energy consumption of nodes in a wireless sensor network with the SAD_ ML technique comparison of results with the literature.

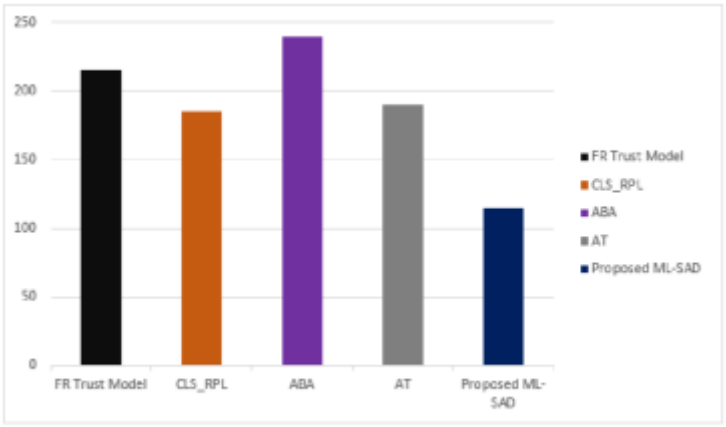

Figure 13. Energy Consumption comparison with literature

Evaluation Metrics

The proposed scheme evaluated for the following matrices:

Accuracy

The detection accuracy of sinkhole attack detection in wireless sensor networks is measured as follows.

$$\text{Accuracy} = \frac{TP+TN}{} \quad [22]$$

$$\frac{}{TP + FP + TN + FN}$$

Where TP = True Positive, FP = False Positive
TN = True Negative, FN = False Negative

Energy Consumption

Total energy consumed nodes of wireless sensor network for sinkhole attack detection with machine learning (SAD_ML) technique.

$$ET = (E \times n_i) + e_i \quad [3]$$

Where

E = the average energy consumption

ni = the number of data transmissions given out by the sink node

ei = the sink node's current energy value

ET = threshhold energy consumption

Energy Consumption = $\sum_{i=1}^{n} TCi + \mathrm{EC}\, j$ [23]

Where

Transmission Energy Consume source nodes(TC *i*)

Energy Consume Destination Nodes (EC j)

Conclusion

This paper reviews a number of techniques, in which detect sinkhole attacks in wireless sensor network. The paper provides a thorough critical and comparative analysis of the literature that covers all of sinkhole attacks detection approaches. On the basis of detection accuracy, energy usage, packet delivery ratio, and delay, the methods were assessed. The gaps in the literature that were found demonstrate the future range of work to be done in identifying sinkhole attacks with high detection accuracy and low energy consumption in wireless sensor network. We propose sinkhole attack detection with machine learning (SAD_ML) technique with less energy consumption and high accuracy scheme for the classification of sinkhole attack detection in the wireless sensor network. First, we find suspicious nodes by using the Ad hoc on Demand Distance Vector protocol (AODV) and ACK method. In this method, when source node send rout request to designation node then designation node reply to source node. If acknowledgement is not received by source node with in time to live that node is suspicious node. Secondly, suspicious nodes classify by using machine learning classification algorithms for sinkhole attack detection in the wireless sensor network. In terms of detection accuracy, the comparison of machine learning classification algorithms reveals that SVM performs better than the other models. Our SAD_ML technique accuracy 96 percentage with SVM algorithm. The detection accuracy of sinkhole attack in a wireless sensor network with SAD_ML technique accuracies comparison of results with the literature. Less energy consumption of nodes in the wireless sensor network by SAD_ML technique. The future work is other network attacks can find with SAD_ML technique.

Acknowledgments: We thank the anonymous reviewers and the editor for their valuable comments, which helped us to improve the quality and presentation of the paper.
Contribution of athors: To this effort, each author made an equal contribution. The manuscript's published version was approved by all writers after they had read it.
Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Prathapchandran, K. and T. Janani, A trust aware security mechanism to detect sinkhole attack in RPL-based IoT environment using random forest–RFTRUST. Computer Networks, 2021. 198: p. 108413.
2. Yadav, H. and M.S. Tak, Detection of Sinkhole Attack in Wireless Sensor Network Using Ad-hoc on-demand Distance Vector.
3. Al-Maslamani, N. and M. Abdallah. Malicious node detection in wireless sensor network using swarm intelligence optimization. in 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). 2020. IEEE.
4. Jamil, A., M.Q. Ali, and M.E.A. Alkhalec, Sinkhole Attack Detection and Avoidance Mechanism for RPL in Wireless Sensor Networks. Annals of Emerging Technologies in Computing (AETiC), 2021. 5(5): p. 94-101.
5. Nithiyanandam, N. and L. Parthiban, An efficient voting based method to detect sink hole in wireless acoustic sensor networks. International Journal of Speech Technology, 2020. 23: p. 343-354.
6. Mehta, A. and J.K. Sandhu. An Algorithmic Framework for Sinkhole Attack Detection and Mitigation in Wireless Sensor Networks. in 2022 2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE). 2022. IEEE.
7. Jatti, A.V. and V. Sonti, Sinkhole Attack Detection and Prevention Using Agent Based Algorithm. Journal of University of Shanghai for Science and Technology, 2021. 23(5): p. 526-544.
8. Nithiyanandam, N. and P. Latha, Artificial bee colony based sinkhole detection in wireless sensor networks. Journal of Ambient Intelligence and Humanized Computing, 2019: p. 1-14.
9. Nwankwo, K.E. Sinkhole attack detection in a wireless sensor networks using enhanced ant colony optimization to improve detection rate. in 2019 2nd International Conference of the IEEE Nigeria Computer Chapter (NigeriaComputConf). 2019. IEEE.
10. Tak, S. and A. Trivedi, Identifying sinkhole Attack in WSN'S Using Secure AODV. Annals of the Romanian Society for Cell Biology, 2021: p. 6915–6929-6915–6929.
11. Mondal, K., et al., Detecting Sinkhole Attacks in IoT-Based Wireless Sensor Networks Using Distance From Base Station. International Journal of Information System Modeling and Design (IJISMD), 2022. 13(6): p. 1-18.
12. Kumar, D. and E.N. Kapoor. Novel Scheme for Mutual Authentication to Isolate Sinkhole Attack in Wireless Sensor Networks. in 2022 International Conference on Engineering and Emerging Technologies (ICEET). 2022. IEEE.
13. An, G.H. and T.H. Cho, Improving Sinkhole Attack Detection Rate through Knowledge-Based Specification Rule for a Sinkhole Attack Intrusion Detection Technique of IoT. International Journal of Computer Networks and Applications (IJCNA), 2022. 9(2).
14. Raj, P.R. and D. Anand. Sink Hole Attack Detection using Two Step Verification Technique in Wireless Sensor Networks. in 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N). 2021. IEEE.

15. Bhale, P., et al. Energy efficient approach to detect sinkhole attack using roving IDS in 6LoWPAN network. in Innovations for Community Services: 20th International Conference, I4CS 2020, Bhubaneswar, India, January 12–14, 2020, Proceedings 20. 2020. Springer.
16. Bilal, A., S.M.N. Hasany, and A.H. Pitafi, Effective modelling of sinkhole detection algorithm for edge-based Internet of Things (IoT) sensing devices. IET Communications, 2022. 16(8): p. 845-855.
17. Abd Halim, I.H., M.H.A. Azziz, and M.F.M. Fuzi, Sinkhole Attack in IDS: Detection and Performance Analysis for Agriculture-based WSN using Cooja Network Simulator. Journal of Computing Research and Innovation, 2021. 6(2): p. 173-181.
18. Karthigadevi, K., S. Balamurali, and M. Venkatesulu. Based on Neighbor Density Estimation Technique to Improve the Quality of Service and to Detect and Prevent the Sinkhole Attack in Wireless Sensor Network. in 2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS). 2019. IEEE.
19. Dhaked, U., A. Kumar, and B.K. Singh, Detection and Isolation Technique for Sinkhole Attack in WSN.
20. Nwankwo, K.E., et al. A Panacea to soft computing approach for Sinkhole attack classification in a wireless sensor networks environment. in Futuristic Trends in Network and Communication Technologies: Third International Conference, FTNCT 2020, Taganrog, Russia, October 14–16, 2020, Revised Selected Papers, Part I 3. 2021. Springer.
21. Keerthika, M. and D. Shanmugapriya, Wireless sensor networks: Active and passive attacks-vulnerabilities and countermeasures. Global Transitions Proceedings, 2021. 2(2): p. 362-367.
22. Salmi, S. and L. Oughdir, Performance evaluation of deep learning techniques for DoS attacks detection in wireless sensor network. Journal of Big Data, 2023. 10(1): p. 1-25.
23. Chen, R.-C., C.-F. Hsieh, and Y.-F. Huang. A new method for intrusion detection on hierarchical wireless sensor networks. in Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication. 2009
24 Kok, S. H., Azween, A., & Jhanjhi, N. Z. (2020). Evaluation metric for crypto-ransomware detection using machine learning. Journal of Information Security and Applications, 55, 102646.
25 Shafiq, M., Ashraf, H., Ullah, A., Masud, M., Azeem, M., Jhanjhi, N. Z., & Humayun, M. (2021). Robust Cluster-Based Routing Protocol for IoT-Assisted Smart Devices in WSN. Computers, Materials & Continua, 67(3).
26 Lim, M., Abdullah, A., Jhanjhi, N. Z., & Supramaniam, M. (2019). Hidden link prediction in criminal networks using the deep reinforcement learning technique. Computers, 8(1), 8.
27 Gouda, W., Sama, N. U., Al-Waakid, G., Humayun, M., & Jhanjhi, N. Z. (2022, June). Detection of skin cancer based on skin lesion images using deep learning. In Healthcare (Vol. 10, No. 7, p. 1183). MDPI.
28 Sennan, S., Somula, R., Luhach, A. K., Deverajan, G. G., Alnumay, W., Jhanjhi, N. Z., ... & Sharma, P. (2021). Energy efficient optimal parent selection based routing protocol for Internet of Things using firefly optimization algorithm. Transactions on Emerging Telecommunications Technologies, 32(8), e4171.
29 Hussain, K., Hussain, S. J., Jhanjhi, N. Z., & Humayun, M. (2019, April). SYN flood attack detection based on bayes estimator (SFADBE) for MANET. In 2019 International Conference on Computer and Information Sciences (ICCIS) (pp. 1-4). IEEE.
30 Zamir, U. B., Masood, H., Jamil, N., Bahadur, A., Munir, M., Tareen, P., ... & Ashraf, H. (2015, July). The relationship between sea surface temperature and chlorophyll-a concentration in Arabian Sea. In Biological Forum–An International Journal (Vol. 7, No. 2, pp. 825-834).
31 Siddiqui, F. J., Ashraf, H., & Ullah, A. (2020). Dual server based security system for multimedia Services in Next Generation Networks. Multimedia Tools and Applications, 79, 7299-7318.
32 Jabeen, T., Jabeen, I., Ashraf, H., Jhanjhi, N., Humayun, M., Masud, M., & Aljahdali, S. (2022). A monte carlo based COVID-19 detection framework for smart healthcare. Computers, Materials, & Continua, 70(2), 2365-2380.

33 Hanif, M., Ashraf, H., Jalil, Z., Jhanjhi, N. Z., Humayun, M., Saeed, S., & Almuhaideb, A. M. (2022). AI-based wormhole attack detection techniques in wireless sensor networks. Electronics, 11(15), 2324.

34 Shahid, H., Ashraf, H., Javed, H., Humayun, M., Jhanjhi, N. Z., & AlZain, M. A. (2021). Energy optimised security against wormhole attack in iot-based wireless sensor networks. Comput. Mater. Contin, 68(2), 1967-81.

35 Wassan, S., Chen, X., Shen, T., Waqar, M., & Jhanjhi, N. Z. (2021). Amazon product sentiment analysis using machine learning techniques. Revista Argentina de Clínica Psicológica, 30(1), 695.

36 Humayun, M., Alsaqer, M. S., & Jhanjhi, N. (2022). Energy optimization for smart cities using iot. Applied Artificial Intelligence, 36(1), 2037255.

37 Ghosh, G., Verma, S., Jhanjhi, N. Z., & Talib, M. N. (2020, December). Secure surveillance system using chaotic image encryption technique. In IOP conference series: materials science and engineering (Vol. 993, No. 1, p. 012062). IOP Publishing

38 Almusaylim, Z. A., Zaman, N., & Jung, L. T. (2018, August). Proposing a data privacy aware protocol for roadside accident video reporting service using 5G in Vehicular Cloud Networks Environment. In 2018 4th International conference on computer and information sciences (ICCOINS) (pp. 1-5). IEEE..