*Computers, Materials & Continua*
DOI:10.32604/cmc.2021.015363
***Article***

Tech Science Press

# Black Hole and Sink Hole Attack Detection in Wireless Body Area Networks

**Rajesh Kumar Dhanaraj[1], Lalitha Krishnasamy[2], Oana Geman[3,*] and Diana Roxana Izdrui[4]**

[1]School of Computing Science and Engineering, Galgotias University, 201310, India
[2]Department of Information Technology, Kongu Engineering College, Erode, 638060, India
[3]Department of Health and Human Development, Stefan cel Mare University, Romania
[4]Department of Computers, Electronics and Automation, Stefan cel Mare University, Romania
[*]Corresponding Author: Oana Geman. Email: oana.geman@usm.ro
Received: 17 November 2020; Accepted: 08 February 2021

**Abstract:** In Wireless Body Area Networks (WBANs) with respect to health care, sensors are positioned inside the body of an individual to transfer sensed data to a central station periodically. The great challenges posed to health-care WBANs are the black hole and sink hole attacks. Data from deployed sensor nodes are attracted by sink hole or black hole nodes while grabbing the shortest path. Identifying this issue is quite a challenging task as a small variation in medicine intake may result in a severe illness. This work proposes a hybrid detection framework for attacks by applying a Proportional Coinciding Score (PCS) and an MK-Means algorithm, which is a well-known machine learning technique used to raise attack detection accuracy and decrease computational difficulties while giving treatments for heartache and respiratory issues. First, the gathered training data feature count is reduced through data pre-processing in the PCS. Second, the pre-processed features are sent to the MK-Means algorithm for training the data and promoting classification. Third, certain attack detection measures given by the intrusion detection system, such as the number of data packages trans-received, are identified by the MK-Means algorithm. This study demonstrates that the MK-Means framework yields a high detection accuracy with a low packet loss rate, low communication overhead, and reduced end-to-end delay in the network and improves the accuracy of biomedical data.

**Keywords:** Wireless body area network; black hole attack; sink hole attack; proportional coinciding score; intrusion detection; correlation rate

## 1 Introduction

One of the most active research areas in the past few years is that regarding wireless sensor networks (WSNs). In WSNs, sensor nodes cooperate with one another by sensing the deployed surroundings and updating to the Sink. Sink hole attack attracts the network traffic in neighborhood nodes by promoting itself to own the shortest route to the Sink.

Based on the existing route, the sink hole node attempts to invite the stream of traffic through a specific area. The rest of the nodes in the network then use this malicious node path and interchange their data. Given that the communication dependency for affected nodes is through a malicious node, the sink hole attack easily creates a way for other kinds of attacks, such as gray hole and black hole. A distributed adaptive framework is proposed [1] on the basis of the subjective opinion and probabilistic logic extension of scheduled automata to identify the likelihood of each sensor node being attacked by sink hole. Fig. 1 shows the model of the sink hole attack.

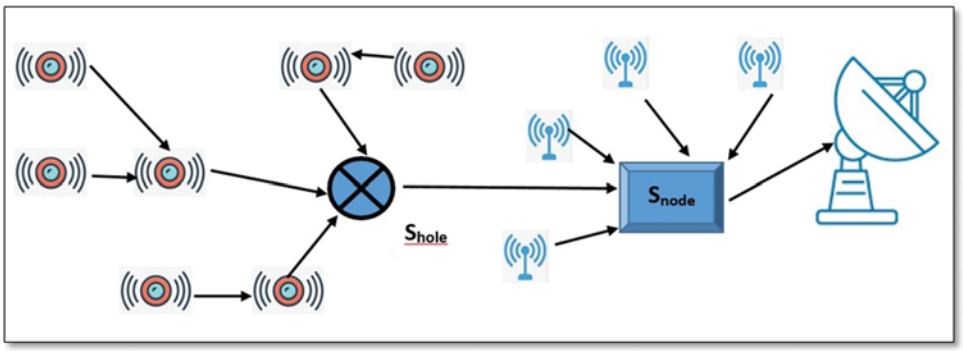


**Figure 1:** Sink hole attack

As presented in Fig. 1, the red sensors represent the compromised nodes, whereas the blue sensors represent the normal sensor nodes with "$S_{hole}$" and "$S_{node}$" being the sink hole and sink node, respectively. In [1], subjective opinion models at the Sink evaluate the probabilistic logic in an iterative manner. This evaluation is conducted according to node behaviors, such as +ve and −ve observations collected by the distributed sensors, which are adaptively tuned on the basis of timed automation. This timed automation captures the entire network behavior at runtime in the Sink. Given that the routing paths are selected from reliable nodes, the packet loss ratio is reduced.

To determine the discriminative features of other attacks in Wireless Body Area Networks (WBANs), such as black hole and worm hole, the overlap between positive and negative observations to be analyzed is an important performance criterion. Malicious nodes utilize the holes involved during the route discovery process with the objective of carrying out their malicious intent. One such familiar attack is the black hole attack, which sends fake routing information to a source node and drops the entire data after hosting itself in the path among source nodes and destinations. A malicious node sends a fake reply to the start node that it possesses the shortest route (i.e., malicious route request) to the end node. The start node establishes a path with this malicious node by sending data packets to it, which, in turn, discards the transmitted data packets. Fig. 2 illustrates a sample scenario of black hole attack.

In Fig. 2, "$S_a$" denotes the start node (i.e., sensors), whereas "$S_e$" denotes the end node with a malicious route reply obtained from sensor "$S_d$" and a normal route reply from "$S_b, S_f, S_c$" with a route request placed by "$S_a$." An effective security algorithm against the black hole attack is given in [2], which identifies the secured path by improving the processing of data packet with minimum end-to-end (E2E) delay and routing overhead. Such an improvement is achieved using a threshold value built on the request path ($P_{REQ}$) sent and reply path ($P_{REP}$) received. However, the existing solutions provided in [3] exclude solutions for other attacks.

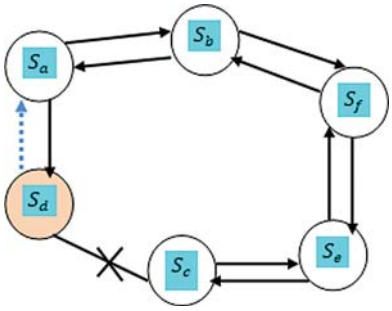


**Figure 2:** Black hole attack

The sink hole attack in MANETs is distinguished using multi-hop routes from the start node to the end node, as investigated in [4]. A unique cooperative cross layer approach is presented in [5] to discover the black hole attack by using multipoint relays for establishing the route [5]. Black hole and sentinel attacks are addressed in [6] on the basis of the decision process, which includes the Bayes rule and a simple threshold approach. The effect of a cooperated node on the zone routing protocol is designed in [7] to mitigate the black hole attack, resulting in the improvement of packet delivery ratio, throughput, and E2E delay. Although improvement is attained, authentication remains a major issue. To resolve this issue, a technique to discover a pattern on the basis of gathered passwords and an attempt to provide a worker's password using hash are depicted in [8].

Although the subjective logic and trust-based model are often used in different networks, the means to address different types of attacks and the overlapping between observations remain unaddressed. The prevention of other types of attacks may not be assured using the subjective logic and trust-based model. This work leverages the flaws of the subjective logic, in addition to the trust-based model and the method to combat against the black hole and sink hole attacks, through an overlapping score and by using machine learning techniques.

This paper is organized as follows: Section 2 includes briefings about existing methods to overcome the different types of attacks. Section 3 explains the sink hole and black hole attack detection methodology. Section 4 presents the parametric explanations and experimental settings. Section 5 summarizes the impacts of different performance metrics by evaluating up-to-date works. Section 6 concludes with limitations and future scope.

## 2 Literature Review

With the wide popularity of next-generation communication networks, such as MANETs and vehicular ad-hoc networks (VANETs), safety has become a major concern. Wireless sensors and delay tolerant network applications have emerged as promising technologies; some of them are used for health care, smart grid, and target monitoring. A geometric-based detection of black hole and gray hole attack schemes is investigated in [9]. In [10], a secure data fragmentation is designed to avoid and identify the sink hole and Sybil attacks on the existence of fixed and dynamic deployed nodes, yielding high detection accuracy and meager false-positive rate. In addition to issues and taxonomy, future directions against DDoS attacks are presented in [11]. A survey of Sybil defense algorithms used in online social webs is provided in [12].

A mutable black hole unearthing mechanism is designed in [13] to identify the behavioral changes of nodes. However, the mechanism fails to identify the black hole attack. WBANs are

prone to many attacks, which weaken network performance. Black hole attack represents the security risk that reduces the normal functioning by dropping all received packets. Even though many mechanisms are compared in [14] to protect the network from the black hole attack, an effective and lightweight security mechanism detects and prevents DDoS attacks.

An attempt is made to investigate the features for ensuring secure routing protocols in VANETs and combat against the black hole attack by applying the stack operations, as given in [15,16]. Another measure to fight against the black hole attack is by using the AODV routing protocol, which is designed in [17]. However, the reliability of the node is not ensured and thus compromises the entire network. A proficient GA-based denial of sleep attack detection is proposed in [18]; it guarantees the reliability of the connector node by using the fitness function estimation.

Node attacks in WBANs are unavoidable, owing to node replication. To address this issue, a novel type named Solo Stage Random Walk Memory with a Distributed Network is proposed in [19]; it guarantees the security of a node with sensible overhead with respect to memory and communication. However, with nodes prone to various attacks, data communication is found insecure in both ways. Although data communication is ensured in [20], attack detection is said to be compromised as network density increases. To address this issue, different types of sink hole nodes are identified in [21] on the basis of several disjoint clusters. An intrusion detection with various machine learning techniques is surveyed and presented in [22,23].

Some secured solutions are provided to protect WBANs from the black hole attack. An improvised hierarchical efficient intrusion detection system (IDS) preserves the sensor network from such an attack [24,25]. The approach depends on the exchange of control packets between the sensor node and base station. However, the security remains unaddressed.

Many security keys are used to secure data transmission from the black hole attack. A hierarchical energy efficient IDS, which preserves the sensor network from the black hole attack, is designed in [26]. The designed system controls the packet exchange between the sensor node and base station. However, the designed system still experiences network overhead issues.

A new probabilistic approach identifies and separates the black hole attack from MANETs. Routing algorithms are not designed to reduce and remove this type of attack. A technique called Novel Honeypot-based Detection and Isolation helps identify and remove the black hole attack. The Honeypot method increases the security rate of MANETs with minimum network overhead. However, such a method is cost-ineffective [26,27]. A new technique identifies the black hole attack by multiple base stations and performs verification through an agent-based technology [28].

From the observations drawn, many of the techniques and/or algorithms are implemented for detecting attacks yet suffer from overlapping discriminative features on the intermediary node, resulting in computational issues in identifying attacks. This work proposes a novel algorithm with the objective of reducing the computational complexity (CC) of identifying the discriminative features by using the proportional coinciding score (PCS).

## 3 Methodology of MK-Means

A scenario for remote patient monitoring is considered in this work, where the PCS is used to collect data and bring alert activation when abnormal patterns (i.e., sink hole or black hole nodes) are detected, as followed in [29,30]. The proposed framework is based on the machine learning technique given in [31,32]. It clusters the sensor nodes according to the updated distance

(as patient records are dynamic) by using the MK-Means algorithm. Fig. 3 depicts the block structure of hybrid, sink hole, and black hole attack detection framework.

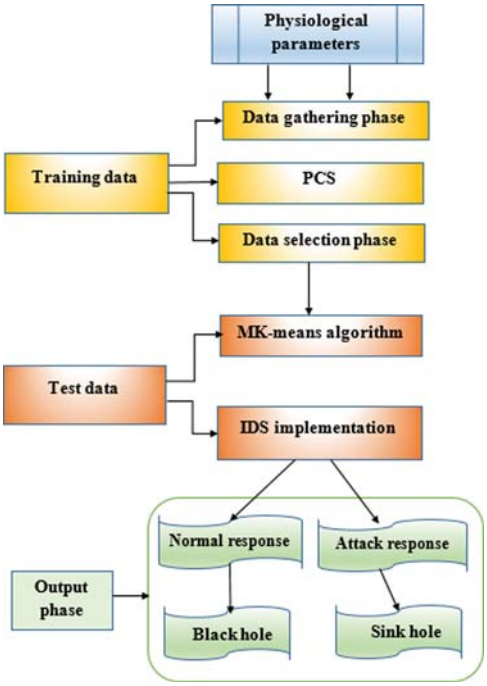


**Figure 3:** Block diagram of MK-Means

The block diagram consists of three different phases, namely, physiological data gathering (PDG) phase, data reduction phase, and classification phase. The description of each module is elaborated below.

### *3.1 PDG*

The PDG phase gathers data or features from the physiological parameter measurements presented in the form of data matrix "$DMat = DMat_{ij}$." In the data matrix, "$i$" represents the time-related growth values, whereas "$j$" represents the measured parameter. All gathered values for all measured parameters are stored as a single record incrementally at time "$t$." It is represented by "$DMat_t = (DMat_{t1}, DMat_{t2}, \ldots, DMat_{tn})$" for each patient and considered packet "$P_t$."

The recorded features are denoted by "$Fe = (Fe_1, Fe_2, \ldots, Fe_n)$," where "$Fe_i$" represents column "$i$" in data matrix "$DMat$." The features represent patient metrics, such as heart rate "$H_{Rate}$" and blood pressure "$B'_P$." Hence, the data matrix is written as given below.

$$\begin{array}{c} \quad Fe_1 \quad Fe_2 \quad Fe_n \\ DMat = \begin{array}{c} DMat_1 \\ DMat_2 \\ \ldots \\ DMat_n \end{array} \begin{bmatrix} dmat_{11} & dmat_{12} & \ldots & dmat_{1n} \\ dmat_{21} & dmat_{22} & \ldots & dmat_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ dmat_{m1} & dmat_{m2} & \ldots & dmat_{mn} \end{bmatrix}. \end{array} \tag{1}$$

The PDG module in the hybrid attack detection framework monitors the events and packets, their delivery time, and topological changes; the module also archives feature inputs. The MK-Means framework selects the hidden data, which tolerate evidence of familiarity or anomaly. The healthcare WSN is in need of the received physiological data to maintain its accuracy. Wrong triggering or erroneous instances when required to alert medical personnel may result in fatal health issues [33]. To increase the attack detection accuracy and reduce false alarms (improper diagnosis), black hole and sink hole attacks are given focus in this study.

### *3.2 PCS*

In the MK-Means framework, an analytical method for opting features, which are based on an overlying examination of medical data through classes for each patient, is presented. In the data or feature selection module, data packets are distributed to the PCS technique to select the best features.

The PCS is determined with the objectives of exploring the commonality between features across classes for unalike patients and identifying discriminative features. This approach employs the facts given by training classes along with testing classes to explore contradictory features among target classes for hybrid attack detection.

Physiological data can generally be presented in the matrix form, "$DMat = dmat_{ij}$," where "$DMat \in K * N$." Here, "$i = 1, 2, \ldots K$" and "$j = 1, 2, \ldots N$." Each sample record is categorized by a labeled target class "$P_j$," signifying the patient model. Class labels represented as vectors, such as "$P \in N$" and its "$jth$" element "$p_j$," have a unique value called "$Cl$," which may be either "1" or "2."

Minimum subsets of patient records (i.e., features) are selected using the PCS method, and the data delivered are analyzed. This resultant value is assumed to be the small one that suitably classifies the huge set of samples in the training set given. This kind of procedure agrees to the discarding of repeated samples, such as patient records with identical profiles, thus avoiding overlapping.

Let "$Fe$" be taken as a set comprising all features (i.e., "$|Fe| = N$"). Consider also "$A_M(F)$" to be the features aggregate mask, which is stated as the rational disconnection corresponding to patient records, which belong to the set. It is expressed as given below.

$$A_M(F) = A_{M1} \vee A_{M2} \vee \ldots A_{MN}. \tag{2}$$

Then, the mathematical formula to cover the maximum number of patient records is as follows:

$$max\left(\sum_{j=1}^{N}\left(dmat_{ij} = 1\right)\right). \tag{3}$$

From Eq. (3), the patient record set whose data matrix values have the maximum bits of "1" is assigned. The objective of this proposed framework should be to look for the tiny subset, indicated by "$Fe'$," by which "$A_M\left(Fe'\right)$" matches to the features aggregate mask in the bunch of patient records and "$A_M(Fe)$" fulfilling the following assertion.

$$min\left(\left|Fe'\right|\left(A_M\left(Fe'\right) = \vee A_{Mi} = A_M(Fe)\right)\right). \tag{4}$$

The above mentioned phenomenon is iterated successively and stops when all patient records are processed, that is, when the selected patient record covers almost all the samples. The step-by-step instruction of minimizing the redundancy and maximizing the relevancy by using the PCS is given in Algorithm 1.

**Algorithm 1:** Minimizing the redundancy and maximizing the relevancy

**Input:** Patients "$N_p$," packet "$Pkt_t$," time "$t$," features "$Fe = (Fe_1, Fe_2, \ldots, Fe_n)$"
**Output:** Minimal redundancy, maximum relevant features
1: **Begin**
2: **For** each packet "$Pkt_t$" with "$N_p$" patients
3: Frame the data matrix with features by using Eq. (1)
4: Find the unique feature from feature set $Fe$ by using Eq. (2) and bring the logical disconnection among features
5: Acquire maximum relevance by using Eq. (3) to cover the maximum number of patient records
6: Minimize the redundancy of features identified in the patient records by using Eq. (4)
7: **End for**
8: **End**

As given in Algorithm 1, the PCS in the MK-Means framework assigns the minimum subset of patient records (i.e., not all the attributes of the records are considered; only their relevant attributes), giving the maximum likelihood of classification accuracy in a selected training set. Patient records with a minimal subset are integrated with highly measured patient records on the basis of the "*PCS*," resulting in an absolute feature selection.

### *3.3 MK-Means Clustering*

K-means clustering is used to identify sets of sensor nodes with their collection represented by variable "$K$." As patient records (e.g., blood pressure, pulse rate, etc.) change with respect to time, the MK-Means algorithm is used in this work because of the advantage of using rescaled (i.e., varied with respect to time) entity points.

K-means clustering involves a set of sensor nodes "$N$," "$i$" set of features, and "$V$;" including data matrix "$DMat = dmat_{ij}$," where "$dmat_{ij}$" is the value of feature "$j \in J$" at entity "i $\in I$." The framework makes a partition "$\{Fe' = Fe'_1, Fe'_2, \ldots, Fe'_n\}$" of "$i$" in "$K$" non-overlapping subsets generated by Algorithm 1, referred to as clusters, each represented by a centroid "$c_K = (c_{kv})$" in the feature space "$k = 1, 2, \ldots, K$." Then, "within cluster distances to centroid" are measured and given below.

$$V\left(Fe', c\right) = \sum_{i=1}^{n} Dis\left(i, c_k\right). \tag{5}$$

Subsequently, the Minkowski m-metric between M-dimensional points "$x = x_h$" and "$y = y_h$" is defined using the equation given below.

$$Dis_m\left(x, y\right) = \sum_{v=1}^{n} |x_h - y_h|^m - v\left(Fe', c\right). \tag{6}$$

With the obtained M-dimensional points, the groups or similar nodes (i.e., similar records), which exist in the network, are identified. With the obtained information, the identified records are transferred to the cluster head (CH) nodes. Fig. 4 illustrates the block diagram with a brief working structure of the suggested IDS by using the MK-Means algorithm. The work proposed here assumes that the sink hole or black hole nodes either drop the data or claim to owe the shortest path.

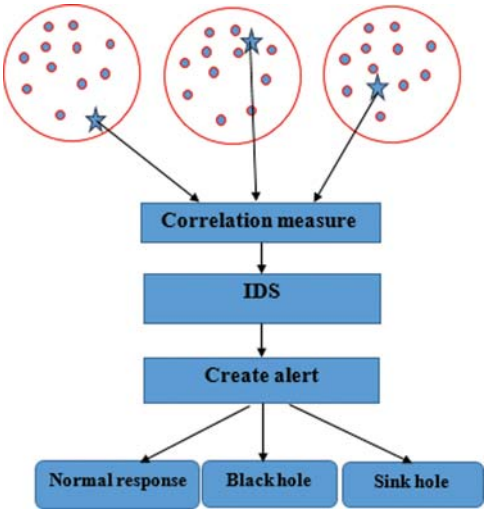


**Figure 4:** MK-means IDS

From the sensing devices, the CH gathers the data, and nodes are differentiated by stars as the CH and normal sensor nodes are denoted by filled circles. The Sink position is fixed randomly on the basis of the GCCR [34].

The intrusion detection consists of a correlation measure that estimates the intrusion measure ($Int_{Msr}$) from the extracted feature values. Data packets receive "$DPkt_{rec}$;" data transmit packet "$DPkt_{tr}$;" and node ID's "$S_i$" is applied to estimate "$Int_{Msr}$." The alert message is activated by an intrusion detection engine that depends on the "$Int_{Msr}$" input, which denotes the presence or absence of cooperated nodes (i.e., sink hole or black hole nodes). Fig. 5 shows an example network that includes sink hole and black hole attacks.

As presented in Fig. 5, "$S_1$" and "$D_{st}$" are the source and destination nodes, respectively. Circle "$S_h$" represents the sink hole node, which attracts the traffic with "$S_4$," "$S_6$," and "$S_7$" being the compromised nodes. Meanwhile, "$S_9$" denotes the black hole node, which claims to possess the shortest path with "$S_5$" being the compromised node.

The correlation measure is obtained with the aid of the intrusion measure using the data packets received and data packets transmitted for each CH node. It is mathematically measured as follows:

$$\sum_{i=1}^{n} Int_{Msri} = \sum_{i=1}^{n} \frac{DPktR_i}{DPktT_i}. \tag{7}$$

With the intrusion measure, the IDS activates an alert message according to the resultant values.

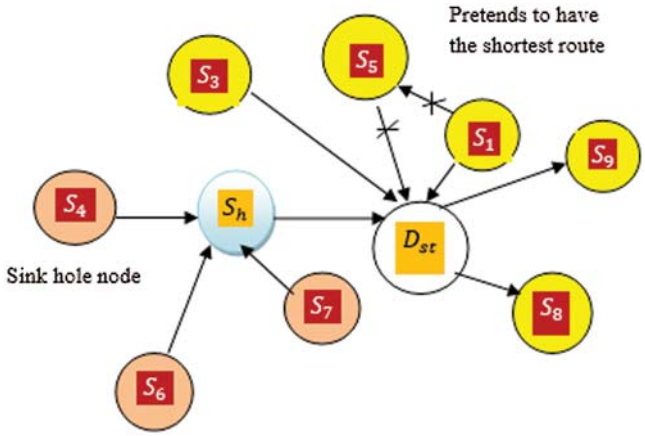


**Figure 5:** An attack detection generated by the IDS with a sample network

The pseudo code representation for the sink hole and black hole attack detection using a machine learning technique is given in Algorithm 2.

**Algorithm 2:** Sink hole and black hole attack detection using a machine learning algorithm

**Input:** Sensor node "$S_i$," data packet received "$DPktR$," data packet transmitted "$DPktT$"
**Output:** Detection of sink hole or black hole

**Begin**
  **For** network sensing nodes "$S_i$" (i.e., that form cluster) $<= N$
    Calculate in cluster distances by using Eq. (5)
    Calculate m-metric between M-dimensional points by using Eq. (6)
  **End for**
  **Repeat**
    Calculate intrusion measure "$Int_{Msr}$" by using Eq. (7)
    **If** "$Int_{Msri} \rightarrow \infty$," **then**
      The matching "$S_i$" is the sink hole node.
      Isolate "$S_i$" node
      Send alert notice about "$S_i$" to the cluster member nodes that remain in the network
      **End if**
      **If** "$freq\,(DPktR_i\,[S_i]) = 0$," then
      Fix black hole node "$S_i$"
      Isolate "$S_i$" node
      Send alert notice about "$S_i$" to the cluster member nodes that remain in the network
    **Else**
      The matching "$S_i$" in the network is a normal CH.

19: **End if**
20: **Until** communication process is accomplished
21: **End**

The IDS comprises a correlation measure module, which calculates the intrusion measure ($Int_{Msr}$) from the extracted features by using the PCS. The correlation measure transmits the "$Int_{Msr}$" input to the attack detection engine. The detection technique activates the alert message depending on the "$Int_{Msr}$," which signifies the role of compromised node as live or not.

## 4 Experimental Results

The model specified is executed using Network Simulator-3 (NS-3). Nodes, which are deployed in the region on the basis of the random-waypoint model and scenario, vary from 50 to 500. For each sensor node, the model behavior is repeated individually, and variation is made in mobility by constructing each node to be stationary for a certain pause period. The parameters used for simulation are shown in Tab. 1.

In this segment, the experimental results are presented to exhibit the efficiency of the hybrid attack detection accuracy framework. The proposed framework is compared along with the adaptive sink aware (SINK-AWARE) algorithm and secure route discovery in AODV (AODV-SR) as both are analogues to the proposed framework; the location estimation of sink hole and black hole attackers is also conducted using the two algorithms. In the NS-3 implementation, a set of nodes in WBANs is randomly deployed in a "$500 * 500$ $m^2$" area, and each node has 150 m as the communication range.

**Table 1:** Parameters used for simulation

| Parameter | Input |
|---|---|
| Simulation round time | 500 seconds |
| Halt duration | 10 seconds |
| Sensor nodes | 50–100 (varies up to 500) |
| Sensor node mobility | 10 m/seconds |
| Routing protocol | Dynamic source routing |
| Mobility deployment model | Random waypoint |
| Number of data packets | 100–200 |
| Packet size | 512 Kbits |

### *4.1 Performance Measures*

The following performance metrics are applied in the current scenario.

#### *4.1.1 Computational Complexity (CC)*

*CC* involves the time taken to measure the Minkowski m-metric between M-dimensional points "$Dis_m(x, y)$" and the intrusion measure ($Int_{Msr}$) for detecting intrusion.

$$CC = time(Dis_m(x, y)) + time(Int_{Msr}). \tag{8}$$

#### 4.1.2 Attack Detection Accuracy

The attack detection accuracy is formulated as follows:

$$AD_{acc} = \frac{(S_i - Att_s)}{S_i}. \tag{9}$$

From Eq. (9), attack detection accuracy "$AD_{acc}$" is attained with the complete sensor nodes "$S_i$," which are obtained from patients "$P$" and attack (sink hole or black hole) nodes "$Att_s$" during data transmission in healthcare WBANs.

#### 4.1.3 Packet Loss Rate

Packet loss is said to occur when certain packets fail to reach the intended destination and network congestion. It is measured as the ratio of packets lost with respect to the packets sent.

$$PLR = \sum_{i=1}^{n} (DPktT_i - DPktR_i). \tag{10}$$

From Eq. (10), packet loss rate "$PLR$" is the difference between the data packet transmitted "$DPktT_i$" and the data packet received "$DPktR_i$" by "$i$" nodes.

#### 4.1.4 E2E Delay

E2E delay refers to the duration consumed for a packet to be transmitted to a network wide from source to destination. It is mathematically evaluated as given below.

$$Delay_{End2End} = \sum_{i=1}^{n} Delay \left[ \frac{DPktT_i}{DPkt_i} \right]. \tag{11}$$

E2E delay "$Delay_{End2End}$" is the measure of total delay occurred for the transmission of data packets "$DPktT_i'$" to the successful delivery of a packet with respect to delay to the data packets in total "$DPkt_i$." It is measured in terms of milliseconds and involves waiting at interface queues, propagation time, transfer time, and retransmission time.

## 5 Discussion

The result analysis of the MK-Means framework is compared with existing SINK-AWARE algorithm [1] and AODV-SR [2].

### 5.1 Computational Complexity (CC)

To conduct experimentation, 500 sensor nodes with varying sizes of data packets sent at different time intervals are selected. With these sensor nodes, the CC involved in attack detection is identified.

Fig. 6 illustrates the computational risk involved in attack detection averaged over 500 random training sensor nodes. The inference drawn from the experimental results clearly shows that it exceeds statistical measures compared with existing techniques SINK-AWARE [1] and AODV-SR [2]. The experimental outcome of the proposed work points out a significant improvement in different deployment scenarios and a reduced CC with the MK-Means. With the deployment of 100 nodes, the CC is observed to be 5.25 and 8.89 milliseconds using SINK-AWARE and 9.23 milliseconds using AODV-SR. The CC involved in attack detection is improved with the

application of the PCS, which considers the overlapping of medical data across classes for different patients and evaluates the discriminative features for attack detection. Furthermore, the PCS regarding the overlapping of features is made on the basis of the time-related growth values with which the aggregated mask value is considered, resulting in the minimum redundant and maximum relevant features to be extracted. Moreover, the CC involved in attack detection is reduced by 31% compared with SINK-AWARE and 10% compared with AODV-SR.

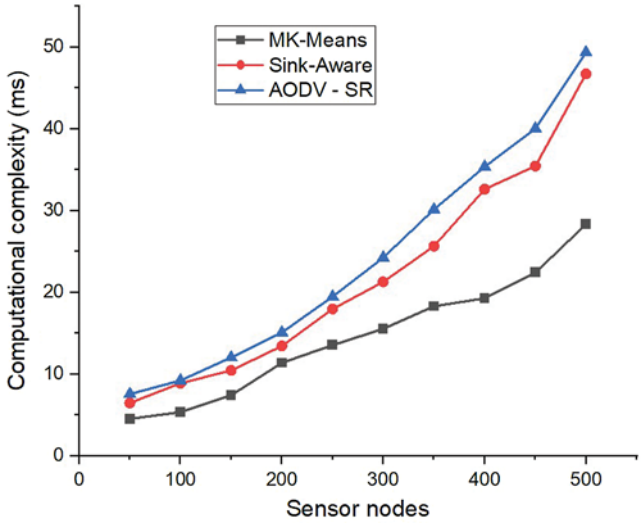


**Figure 6:** Computational complexity with respect to network size

### *5.2 Impact of Attack Detection Accuracy*

The average results of 10 simulation runs, which are performed to measure the accuracy of attack detection, are shown in Fig. 7. The sensing nodes vary from 50 to 500 in the experimental scenario. As discussed in previous sections, the attack (sink hole and black hole) detection ratio using the MK-Means framework produces better results than employing contemporary methods. The attack detection accuracy of the deployment of 50 nodes using MK-Means is 83.93% and 78.53% using SINK-AWARE and 73.53% using AODV-SR.

The attack detection ratio with the MK-Means framework is compared with SINK-AWARE and AODV-SR in Fig. 7, which shows betterment using MK-Means. The MK-Means framework differs from SINK-AWARE and AODV-SR, as it incorporates a machine learning algorithm for sink hole and black hole detection. The advantage of applying a machine learning algorithm in the MK-Means framework is that an intrusion measure is used to validate attacks, and the intrusion measure value is a time-related growth value in that it obtains patient information with respect to time. This time-related increase value is then used to decide whether any type of attack or a normal node is observed. This decision, in turn, improves the attack detection ratio as 7.8% against SINK-AWARE and 8.5% against AODV-SR.

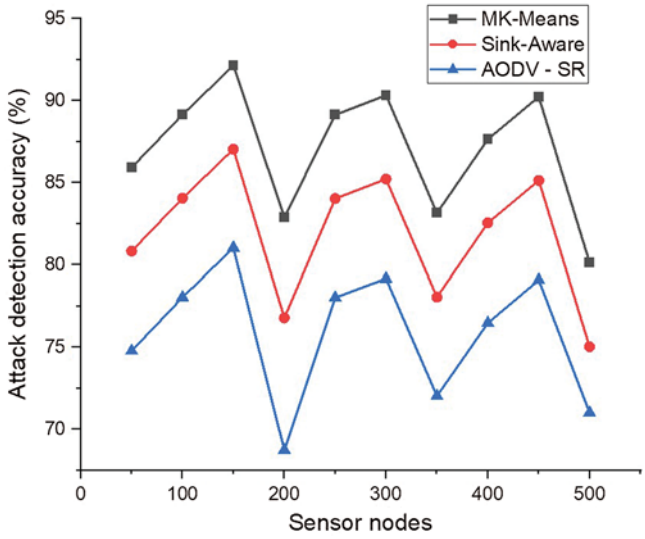


**Figure 7:** Detection accuracy based on the number of nodes

### *5.3 Effect of Packet Loss Ratio*

The packet loss ratio is evaluated when the source transfers a volume of data (i.e., patient records) to the destination (i.e., laboratory technicians or doctors). The experiments are conducted using 200 data packets with various sizes, and the packet loss rate is measured by the set of packets received per second (pps).

Fig. 8 illustrates the packet loss rate for the MK-Means framework, SINK-AWARE, and AODV-SR with 200 different data packets. The average packet loss rate returned over the MK-Means framework increases gradually for different data packets and proves to be efficient when compared with the two other methods. Fig. 8 also reveals that the average packet loss ratio while performing attack detection is enhanced by the MK-Means framework. With the MK-Means algorithm, rescaled entity points with various patient features are considered (i.e., due to the non-static nature of patient records as discussed in [35]). The packet loss rate is reduced by considering the rescaled entity points and non-overlapping subsets produced by Algorithm 1. From the extracted features, attack detection is performed, resulting in the upscale of the packet loss ratio by the MK-Means structure by 25.4% with respect to SINK-AWARE and 16.3% with respect to AODV-SR.

### *5.4 Impact of E2E Delay*

The E2E delay is measured with different sizes of data packets against all the three algorithms, and the resultant values are shown in Fig. 9. The E2E delay rate is measured in terms of milliseconds for experimental purpose. Fig. 9 reveals that the E2E delay is lower when using the MK-Means framework than when using existing methods, such as SINK-AWARE and AODV-SR. The reason is because within cluster distances to centroid are measured using MK-Means, where similar patient records are transferred to the CH node, as given in [36]. This CH node then obtains the correlation by using the intrusion measure. Therefore, not all the node information is transferred to the IDS; only the updated patient records with respect to time using MK-Means are used as measures for obtaining the correlation value. Considering such a situation, the E2E

delay, which uses the MK-Means framework, is reduced by 37.4% compared with SINK-AWARE and 8.7% compared with AODV-SR.

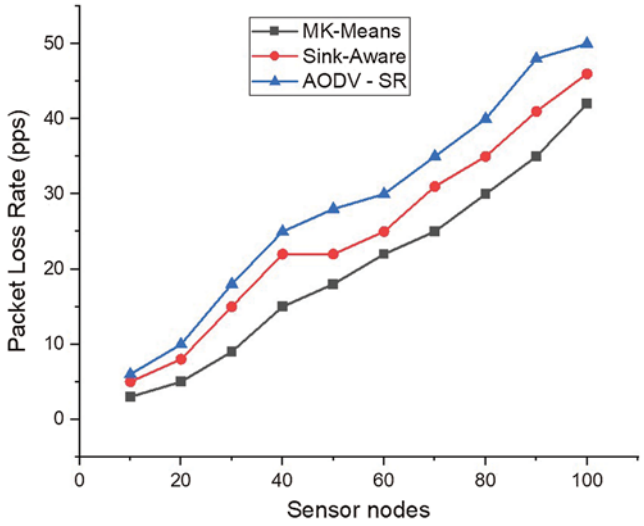


**Figure 8:** Packet loss rate

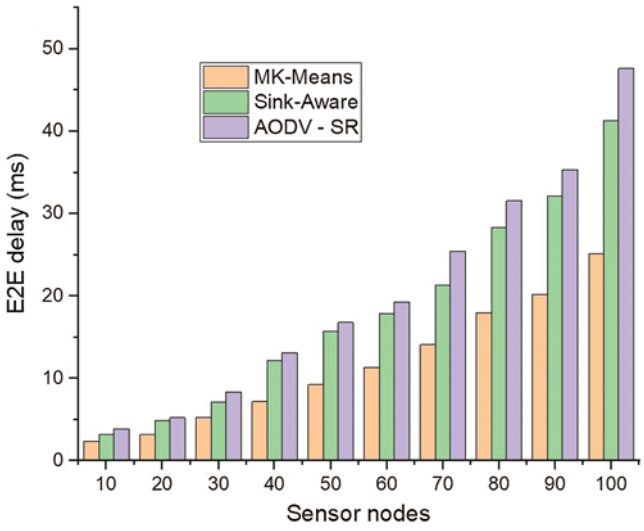


**Figure 9:** E2E delay with increased sensor nodes

## 6 Conclusion

Healthcare WBANs remain susceptible to security risks, such as sink hole and black hole attacks. To overcome these issues, an intrusion detection framework is proposed for detecting attacks and alerting sensing nodes to minimize data loss. Specifically, the architectural view for the intrusion detection of healthcare WBANs, which use the PCS and the MK-Means machine learning technique, is implemented for data reduction and classification accuracy, respectively. The

PCS minimizes the feature size, which, in turn, cuts down classification complexities, with which the attack detection accuracy is improved. The proposed attack detection framework captures the sink hole and black hole nodes with computation ability and activates the remaining deployed nodes through an alert message. Furthermore, the numerical results drawn from various executions demonstrate that the proposed attack detection framework reduces the E2E delay and packet loss rate with the help of CH nodes, which send only updated patient records.

Although CC is reduced, the experiments are conducted with only few records. For handling voluminous data, algorithmic techniques must be changed accordingly or suitable learning algorithms with cloud data management should be incorporated in the future to adopt this technique in healthcare domains.

**Funding Statement:** This research received no grant funding. APC was funded by Ştefan cel Mare University of Suceava, Romania.

**Conflicts of Interest:** The authors declare no conflicts of interest to report regarding the study.

## References

[1] G. Jahandoust and F. Ghassemi, "An adaptive sinkhole aware algorithm in wireless sensor networks," *Ad Hoc Networks*, vol. 59, pp. 24–34, 2017.

[2] J. Kumar, M. Kulkarni, D. Gupta and S. Indu, "Secure route discovery in AODV in presence of black hole attack," *CSI Transactions on ICT*, vol. 3, no. 2–4, pp. 91–98, 2015.

[3] O. Mahmoud, A. Harrison, A. Perperoglou, A. Gul, Z. Khan *et al.,* "A feature selection method for classification within functional genomics experiments based on the proportional overlapping score," *BMC Bioinformatics*, vol. 15, no. 1, pp. 274, 2014. https://doi.org/10.1186/1471-2105-15-274.

[4] L. Sánchez-Casado, G. Maciá-Fernández, P. García-Teodoro and N. Aschenbruck, "Identification of contamination zones for sinkhole detection in MANETs," *Journal of Network and Computer Applications*, vol. 54, pp. 62–77, 2015.

[5] R. Baiad, O. Alhussein, H. Otrok and S. Muhaidat, "Novel cross layer detection schemes to detect blackhole attack against QoS-OLSR protocol in VANET," *Vehicular Communications*, vol. 5, pp. 9–17, 2016.

[6] A. Giaretta, S. Balasubramaniam and M. Conti, "Security vulnerabilities and countermeasures for target localization in bio-nanothings communication networks," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 4, pp. 665–676, 2016.

[7] K. Purohit, S. Dimri and S. Jasola, "Mitigation and performance analysis of routing protocols under black-hole attack in vehicular ad-hoc network (VANET)," *Wireless Personal Communications*, vol. 97, no. 4, pp. 5099–5114, 2017.

[8] H. Mun and K. Han, "Blackhole attack: User identity and password seize attack using honeypot," *Journal of Computer Virology and Hacking Techniques*, vol. 12, no. 3, pp. 185–190, 2016.

[9] T. Pham and C. Yeo, "Detecting colluding blackhole and greyhole attacks in delay tolerant networks," *IEEE Transactions on Mobile Computing*, vol. 15, no. 5, pp. 1116–1129, 2016.

[10] A. Razaque and S. Rizvi, "Secure data aggregation using access control and authentication for wireless sensor networks," *Computers & Security*, vol. 70, pp. 532–545, 2017.

[11] G. Somani, M. Gaur, D. Sanghi, M. Conti and R. Buyya, "DDoS attacks in cloud computing: issues, taxonomy, and future directions," *Computer Communications*, vol. 107, pp. 30–48, 2017.

[12] M. Al-Qurishi, M. Al-Rakhami, A. Alamri, M. Alrubaian, S. Rahman *et al.,* "Sybil defense techniques in online social networks: A survey," *IEEE Access*, vol. 5, pp. 1200–1219, 2017.

[13] S. Malhotra and S. Kumar, "Detection of black hole in ad-hoc networks," *International Journal of Computer Applications Technology and Research*, vol. 3, no. 6, pp. 374–376, 2014.

[14] C. Lal and A. Shrivastava, "An energy preserving detection mechanism for blackhole attack in wireless sensor networks," *International Journal of Computer Applications*, vol. 115, no. 16, pp. 32–37, 2015.

[15] P. Tyagi and D. Dembla, "Performance analysis and implementation of proposed mechanism for detection and prevention of security attacks in routing protocols of vehicular ad-hoc network (VANET)," *Egyptian Informatics Journal*, vol. 18, no. 2, pp. 133–139, 2017.

[16] A. Kumar, M. Albreem, M. Gupta, M. Alsharif and S. Kim, "Future 5G network based smart hospitals: Hybrid detection technique for latency improvement," *IEEE Access*, vol. 8, pp. 153240–153249, 2020.

[17] U. Gopal and K. Subramanian, "A secure cross-layer AODV routing method to detect and isolate (SCLARDI) black hole attacks for MANET," *Turkish Journal of Electrical Engineering & Computer Sciences*, vol. 25, pp. 2761–2769, 2017.

[18] M. Gunasekaran and S. Periakaruppai, "GA-DoSLD: Genetic algorithm-based denial-of-sleep attack detection in wsn," *Security and Communication Networks*, vol. 2017, pp. 1–10, 2017.

[19] M. Aalsalem, W. Khan, N. Saad, M. Hossain, M. Atiquzzaman *et al.,* "A new random walk for replica detection in wsns," *Plos One*, vol. 11, no. 7, pp. e0158072, 2016.

[20] G. Arulkumaran and R. Gnanamurthy, "Fuzzy trust approach for detecting black hole attack in mobile ad-hoc network," *Mobile Networks and Applications*, vol. 24, no. 2, pp. 386–393, 2017.

[21] O. Osanaiye, A. Alfa and G. Hancke, "A statistical approach to detect jamming attacks in wireless sensor networks," *Sensors*, vol. 18, no. 6, pp. 1691, 2018.

[22] A. Buczak and E. Guven, "A survey of data mining and machine learning methods for cyber security intrusion detection," *IEEE Communications Surveys & Tutorials*, vol. 18, no. 2, pp. 1153–1176, 2016.

[23] M. Rajesh Babu and G. Usha, "A novel honeypot-based detection and isolation approach (NHBADI) to detect and isolate black hole attacks in MANET," *Wireless Personal Communications*, vol. 90, no. 2, pp. 831–845, 2016.

[24] O. Adil Mahdi, Y. Bahar Al-Mayouf, A. Basil Ghazi, M. Abed Mohammed, A. Abdul Wahab *et al.,* "An energy-aware and load-balancing routing scheme for wireless sensor networks," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 12, no. 3, pp. 1312, 2018.

[25] K. Abdulkareem, M. Mohammed, S. Gunasekaran, M. Al-Mhiqani, A. Mutlag *et al.,* "A review of fog computing and machine learning: concepts, applications, challenges, and open issues," *IEEE Access*, vol. 7, pp. 153123–153140, 2019.

[26] U. Khan, S. Agrawal and S. Silakari, "Detection of malicious nodes (DMN) in vehicular ad-hoc networks," *Procedia Computer Science*, vol. 46, pp. 965–972, 2015.

[27] V. Kumar and A. Singhrova, "Movement models-based performance comparison of routing protocols in delay tolerant networks," *International Journal of Innovative Research in Science, Engineering and Technology*, vol. 3, no. 8, pp. 15266–15272, 2014.

[28] C. Ding, L. Yang and M. Wu, "Localization-free detection of replica node attacks in wireless sensor networks using similarity estimation with group deployment knowledge," *Sensors*, vol. 17, no. 12, pp. 160, 2017.

[29] R. Dhanaraj, A. Shanmugam, C. Palanisamy and A. Natarajan, "Optimal clone attack detection model using an energy-efficient GSA-based simulated annealing in wireless sensor networks," *Asian Journal of Research in Social Sciences and Humanities*, vol. 6, no. 11, pp. 201, 2016.

[30] Z. Banković, D. Fraga, J. Moya and J. Vallejo, "Detecting unknown attacks in wireless sensor networks that contain mobile nodes," *Sensors*, vol. 12, no. 8, pp. 10834–10850, 2012.

[31] B. Khalaf, S. Mostafa, A. Mustapha, M. Mohammed and W. Abduallah, "Comprehensive review of artificial intelligence and statistical approaches in distributed denial of service attack and defense methods," *IEEE Access*, vol. 7, pp. 51691–51713, 2019.

[32] L. Krishnasamy, R. Dhanaraj, D. Ganesh Gopal, T. Reddy Gadekallu, M. Aboudaif *et al.,* "A heuristic angular clustering framework for secured statistical data aggregation in sensor networks," *Sensors*, vol. 20, no. 17, pp. 4937, 2020.

[33] M. Wazid, A. Das, S. Kumari and M. Khan, "Design of sinkhole node detection mechanism for hierarchical wireless sensor networks," *Security and Communication Networks*, vol. 9, no. 17, pp. 4596–4614, 2016.

[34] K. Lalitha, R. Thangarajan, S. Udgata, C. Poongodi and A. Sahu, "GCCR: An efficient grid based clustering and combinational routing in wireless sensor networks," *Wireless Personal Communications*, vol. 97, no. 1, pp. 1075–1095, 2017.

[35] S. Mostafa, A. Mustapha, A. Hazeem, S. Khaleefah and M. Mohammed, "An agent-based inference engine for efficient and reliable automated car failure diagnosis assistance," *IEEE Access*, vol. 6, pp. 8322–8331, 2018.

[36] F. Al-Dhief, N. Abdul Latiff, N. Noordini Malik, N. Salim, M. Mat Baki *et al.,* "A survey of voice pathology surveillance systems based on internet of things and machine learning algorithms," *IEEE Access*, vol. 8, pp. 64514–64533, 2020.