International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. #, No.# (Mon-20..)

K-means clustering -based Trust (KmeansT) evaluation mechanism for detecting Blackhole attacks in IoT environment

Shameer M[1], L. Gnanaprasanambikai[2]

[1] Computer Science, Karpagam Academy of Higher Education, Coimbatore, India
[2] Computer Science, Karpagam Academy of Higher Education, Coimbatore, India

E-mail address: mdsameer09@gmail.com, gnanaprasanambikai.lakshmanan@kahedu.edu.in

Received ## Mon. 20##, Revised ## Mon. 20##, Accepted ## Mon. 20##, Published ## Mon. 20##

Abstract: The Internet of Things (IoT) is offering numerous applications and making our lives become easier and more comfortable. However, the significant features lead to various research challenges among security is a main concern as we deal with sensitive information in the IoT environment. The environment opens a loophole for various attacks and those attacks harm the network intentionally. Blackholes are one kind of attack that harms routing operations by dropping all incoming packets. To address this issue, a K-means clustering – based Trust (KmeansT) evaluation mechanism has been proposed. Here, the trust evaluation will be done with the help of both direct observations and recommendations for trust will be given by others. Followed by k means clustering algorithm has been applied to enhance the evaluation mechanism. The blackhole attacks are effectively identified by the proposed model. Mathematical models of the proposed work witness the effectiveness of detection. Simulation results will be analysed by comparing them with the existing similar models in terms of various performance metrics.

Keywords: Internet of Things, Security, Blackhole attack, Trust and K-means clustering

1. Introduction

In the modern era, people expect efficient, robust, and sophisticated operational services in their lives. Consequently, information and communication technology plays a major role in satisfying consumer needs. The Internet of Things is one such prominent and trending technology that helps all aspects of human life. It enhances the values of the business, upgrades customer services, and develops decision-making [1]. It is defined as the arrangement of interrelated digital devices, electrical and mechanical devices, computing devices, network devices, people, animals, and surrounding objects those are having with unique identification and are capable of transmitting information over a communication network. More simply it is defined as the collection of sensor-embedded devices that can capable to communicate with each other. Beyond that those devices can sense the outside environment and do some action based on the data being collected from the external environment [2].

Consequently, IoT offers various applications across various fields including smart agriculture, smart home, smart transport, smart city, smart healthcare, Industrial IoT, smart personal assistance etc [3], [4]. Therefore, the applications range from personal use to industry. More importantly, IoT devices collect real-time data and those data can be processed with the help of Big data analytics so that it is helpful in decision making. In addition to that, Artificial Intelligence also takes part in the working environment of IoT to provide a better user experience. The last two decades have experienced a steady rise in the production and deployment of sensing-andconnectivity-enabled electronic devices, replacing ‘‘regular" physical objects. The resulting Internetof-Things (IoT) will soon become indispensable for many application domains. Smart objects are continuously being integrated within factories, cities, buildings, health institutions, and private homes. Approximately 30 years after the birth of IoT, society is confronted with significant challenges regarding IoT security. Due to the interconnectivity and ubiquitous use of IoT devices, cyberattacks have widespread impacts

E-mail: mdsameer09@gmail.com

on multiple stakeholders. Past events show that the IoT domain holds various vulnerabilities, exploited to generate physical, economic, and health damage. Despite many of these threats, manufacturers struggle to secure IoT devices properly [5].

Though it offers various applications, the significant characteristics such as its resource-constrained nature including limited memory, limited battery power, limited processing power, limited bandwidth, open and shared wireless environment, lack of physical protection, self-organized nature, etc. lead to various research avenues. Therefore, the following open issues are getting attention from the research community. The issues are security, privacy, transport protocol, standardization issues, mobility issues, data integrity, authentication, scalability, energy management, and Quality of Services. Among the research challenges, providing security in IoT is a challenging task hence it is getting much attention among the researchers. The reason is limited processing, storage, and battery capabilities of IoT open a gateway for various attacks. More specifically, the heterogeneous nature of IoT devices creates problem in interoperability that leads to security violations. The entire security issues of IoT can be classified into three major categories. Besides, the security violations happened in almost all the layers of the IoT environment. The following figure depicts the security issues of IoT [6–8].

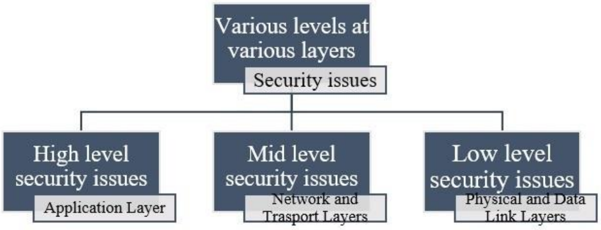
Various levels at various layers
Security issues
High level security issues
Application Layer
Mid level security issues
Network and Trasport Layers
Low level security issues
Physical and Data Link Layers

Figure 1 - Security issues at various layers

The security issues broadly classified into three major categories such as high level, mid-level and low level security issues. The high level security issues are occurred in application layer, mid-level security issues are occurred in both transport and network layers and low level security issues occurred in both network and transport layers. Therefore, to address these attacks several security mechanisms have been proposed by various researchers. Many algorithms like key management, intrusion detection system, blockchain technology, symmetric and asymmetric cryptography algorithms, hash function and etc are effective in providing and ensuring the security in network layer. However, they might not be applicable to resource constrained IoT devices. Applying all those algorithms in the resource constrained IoT devices, leads to security violations [6–8].

The proposed research work is focusing on network layer issues. Issues are raised in the form of attacks and it is defined as an assaulting the system or network environment. Session establishment, RPL routing protocol, insecure neighbour discovery, duplication or replay attack, worm hole attack, blackhole attack, sinkhole attack, Blackhole attack, Sybil attack and buffer reservation attacks are affecting the network layer commonly. The proposed research work is focusing on black hole attack. It is a kind of attack that are affecting the normal routing operation by holding all the incoming packets that are dedicated to forward to others and by the way those nodes are trying to save their energy levels. Result is overall performance of the network become degrade. The main concern of network layer is routing. In IoT environment, data or control packets which are being transmitted from one point to another point with the help of routing protocols. This operation is called routing. Such routing operations are affected by blackhole attacks. Therefore, the entire network operations might be in trouble [9].

To ensure the security the following security requirements must be considered such as confidentiality, authentication, authorization, access control and non-repudiation. Among the requirements authentication considered as primary requirement in the security aspects of IoT environment as it ensures initial level of security. However, in the IoT environment ensuring authentication is complicated task as IoT has heterogeneous devices, cross platform capabilities and resource constrained nature. Therefore, IoT environment is expecting proper authentication along with secure mechanism to protect the IoT environment [10].

The rest of the paper is organized as follows; section 2 deals with the background information which includes k means clustering algorithm, the impact of blackhole attack over RPL routing protocol and prior the working nature of RPL will be discussed, section 3 represents SLR, section 4 unfolds the proposed model, section 5 depicts the algorithm proposed, section 6 deals with the results and discussion and final section deals with the conclusion.

2. Background

The proposed algorithm makes use of K-means clustering algorithm to identify the black hole attack over RPL routing protocol. The following section discusses the k-means clustering algorithm.

A. K means clustering algorithm

It easiest one and it has less computational overhead [11]. This simple algorithm used to categorize the data into K clusters. These clusters have depicted by the centroids [12]. The recognized K-means group is a conventional of data points that are adjoining to the convinced centroid and left since all additional centroids. This algorithm takes approximately deviations. The popularly recycled algorithm is Lloyd's algorithm. In this algorithm, the 'k' amount of bunches has been designated as input with a collection of information points [13].

The procedure starts through starting K-cluster centres. These canters were chosen randomly or based on some heuristic procedure [14]. The center has called the prototype point (centroids). The data points after the data set has allotted to individually cluster based on the closest prototype point. Then, mean data points has calculated by taking the average of data point's coordinate values for separately collection. The mean arguments comprise of a new set of prototype points. Again, every data point has allotted to a cluster of its closest prototype point. This phase of the group is conclusive clustering results. The Euclidean distance has used for proximity measure in K-means. This algorithm consists of many advantages that variety it very familiar. The significant one is simplicity and easy implementation. Since of the direct difficulty, this algorithm mechanism K-Means is a well-known clustering method. It is an unsupervised ML technique to categorize the contribution data groups hooked on numerous modules constructed on Euclidean distance. It remains an algorithm and initiates with the original model opinions [15]. The Euclidean distance is defined as follows:

$$d(x, y) = \sum_{i=1}^{n}(x_i - y_i)^2 \qquad (1)$$

B. The impact of Blackhole over RPL

As discussed in the introduction section, the routing protocols could be used to routing the information from one place to another place. In IoT many routing protocols have been using, however Routing protocol for Low Power Lossy Network (RPL) is often using in IoT environment. The proposed model has also embedded on RPL routing protocol. The detailed discussion on RPL routing protocol has in [16]. The black hole attack is one kind of attack that harm the routing operation by dropping all the incoming packets that are dedicated to forward to others. The figures which has shown below represents influence of black hole attack over RPL.

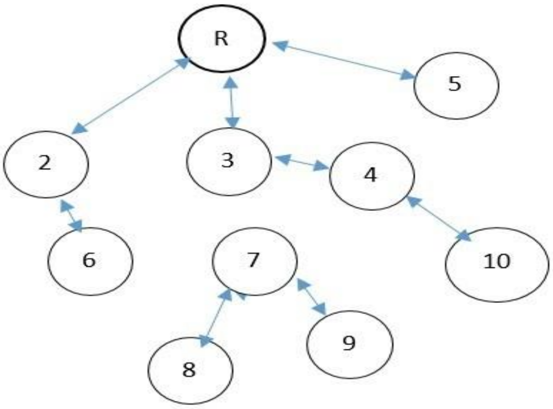

Figure 2 - No black hole attack with RPL

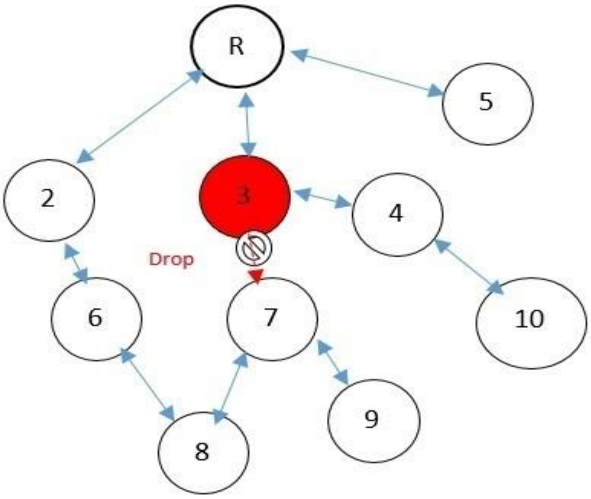

Figure 3 - Black hole attack with RPL

In RPL based IoT environment, initially the devices are authenticated and trusted hence DODAG construction has been done without any difficult. Over the period of time, the behavior of devices might be changed and perform malicious activities like black hole attack. The above "Fig. 3" represent the typical RPL network and RPL network with black hole attack respectively. In the "Fig. 3", an IoT environment is consists of 10 nodes along with the origin node. Here, the node/device 3 has assumed as black hole device. Hence, it publicise the situation that is taking the direct pathway to reach the root device R. therefore, device 7 assuming that device 3 has a path to the root node and forward its packet to that node. As device 3 is block hole node, it will not forward the packet to root node R, by the way it is reducing all the inward packages that are envisioned to onward. By the way black hole attack is executed in the IoBT network.

3. Review of Literature

At present IoT is a hot research topic because of its sustainable development and adoption. The features of IoT leads to various applications and this is also open a gateway for various research avenues in terms of scalability, energy management, security, privacy, interoperability and etc. The security of IoT is getting attention more compare with other research issues as it is threaten to the entire operation of the IoT environment. The following section will discuss some of the existing research that related to security. Several methods and mechanisms have been proposed by various researchers here some of the notable works are pointed out.

The authors in [17] suggested a RPL based routing protocol is called SRAP. It ensures more productivity in non-homogeneous IoT network. It also addressed the scalability issues by providing limited overhead. This model make use of Destination Advertisement Object

(DAO) in encrypted manner to avoid the malicious devices in the network. The authors in [18] suggested a security model which is grounded on Rivest Shamir Adleman public key cryptography algorithm. This model ensure the following security requirements over the RPL based routing protocol such as confidential, integrity and authentication.

The authors in [19] proposed an extendable, secured group, lightweight authentication protocol for IoT environment. The protocol is based on threshold cryptography. This protocol is mainly used to ensure the group communication. The authors in [20] suggested a model to address the attack called blackhole with the help of exponential smoothing algorithm. This is mainly developed to sort out the issues of topological separation executed by the black hole nodes. This method calculates the packet delivery time from the root node. Based on this time, the algorithm make decision and eradicate the black hole attack from the network.

The authors in [21] offered security model which is grounded on Elliptic Curve Diffie- Hellman cryptography. The model ensure the following security properties such as confidentially, authentication, ambiguously, location privacy issue and data packet forwarding security. The authors in [22] suggested a trust prototypical which is built on fuzzy logic. The objective of this trust and fuzzy model is to eliminate black hole attack from IoT by the way trusted route formation. The authors in [23] model is developed to address the issues of both gray hole and warm hole attacks. The trust is designed built on the forwarding and ranking checks. The authors in [24] suggested a trust model which is built on energy. The trust evaluation will be done built on local trust design and collecting the opinions from parent node about their child nodes.

The authors in [25] proposed an authentication mechanism is called Trust Anchor Interconnection loop. It works against the replay and rank spoofing attacks. This model is make use of hash function along with 16 bit unsigned integer to ensure the authentication. As this model is adopted the RPL routing protocol, the parent node will assess the trustworthiness of child nodes in term of authentication.

The authors in [26] suggested lightweight cuckoo filter based security mechanism to address the backhole attack. This model is involved with three stages such as secure rank calculation, infrastructure establishment and node registration. The authentication will be evaluated by border node in the RPL routing protocol. If the nodes are authenticated, those nodes only will be permitted to involve in any network operations. Otherwise, the nodes are not permitted. The authors in [27] suggested a model to mitigate the rank and Sybil attacks. The trust worthiness of devices are designed based on the dependability and reliability of nodes. Besides, positive acknowledgement is also plays a major role in trust computation. The suggested trust model comprises of the following phases: Trust backup, trust computation, monitoring of trust values, identify and separate the malicious nodes and rating the trust. The fuzzy threshold mechanism is used to broadcast the calculated trust values across the networks.

The authors in [28] developed a model to address the Sybil and rank attacks in the IoT environment. The mechanism is called context awareness is used to assess the reliability of both child and parent node in the RPL based IoT network. It is also involved with computation of both direct and indirect trust values. The direct trust is involved with energy, rank value, hop count, unselfishness, and node behavior. Indirect trust is usually from recommendation from parent node as well as child nodes. The authors of [29] suggested a security mechanism based on block chain technology and machine learning algorithm based intrusion detection. The key aim of this suggested model to eliminate internal attacks.

The security mechanisms which are discussed above are effective but till there is a need for effective mechanism as they have drawbacks. Most of the security mechanisms are cryptographic based since it might not be suitable for resource constrained IoT devices. Moreover, they mainly focusing on general security issues not specific to a particular attacks. The essential to counter the above-mentioned issues demonstrated by the prevailing routing safekeeping approaches and the novel dimension of trust organization is the fundamental and likelihood of this proposed work.

4. Proposed Model: K-means clustering algorithm-based trust evaluation (KmeansT)

The aim of the proposed model is to identify and eliminate black hole attacks. To do this, the proposed model follows the underlying assumptions.

A. Assumptions:

- The network environment consists of N number of IoT nodes and they can communicate with each other to do network activities. Then, they can be able to communicate only within their communication range.
- All the participating devices in the environment are resource constrained in terms of their energy, memory and processing capabilities.
- Network consists of fewer number of blackhole nodes to assess the detection capabilities of the proposed model.
- The black hole nodes are called adversaries or compromised nodes and they will not forward the

packets to other nodes and they will drop all the incoming packets by the way it tries to save their energy.

- Every node maintains a trust table where all the trust related information of the participating nodes can be stored. The following table 1 depicts the structure of the trust table.
- The proposed model will execute over a period of time or when the performance of the model decreases.

TABLE I

Nodes' ID	DT	RT	Behaviour

a. Trust Table

B. *Direct Trust*

Over a period of time, the performance of the network may be degraded. In that situation every node in a situation assesses the trustworthiness of its communicating nodes by executing a recommendation-based trust evaluation model. This model consists of the following phases:

- Direct trust evaluation
- Indirect trust evaluation

Classification of blackhole nodes using the K-means clustering algorithm

C. *Direct trust evaluation*

A node Ai wants to evaluate the direct trust of node Aj with the help of the following two factors such as packet forwarding behaviour and confidence level. The packet forwarding behaviour represents how a node can be able to correctly forward the packets to the destination. If the node becomes an adversary node or black hole node, they will not forward the packets as they try to save their energy. Within this consideration, the following equation is used to calculate packet forwarding behaviour.

D. *Packet Forwarding behaviour*

Node Ai wants to calculate the packet forwarding behaviour of node Aj over a period of time. The following equation 2 is used to calculate packet forwarding behaviour.

$$PFB_{Aj}^{Ai}(T) = \frac{P\_F_{(T)}}{P\_D_{(T)} + P\_F_{(T)}} \qquad (2)$$

Where,

$PFB_{Aj}^{Ai}(T)$ denotes packet forwarding behaviour of node Ai with respect to node Aj

$P\_F_{(T)}$ denoted packet forwarding ratio of node Ai

$P\_D_{(T)}$ denotes packet dropping ratio of node Ai

T denotes the time

During the transmission, the blackhole nature of IoT devices makes them become selfish. In that case, those nodes will not perform or be involved in any network operations. This is important to analyse the behaviour of the nodes over time. If the mobile nodes really perform well that can be represented by praise factor β otherwise it is not performing well that can be represented by penalty factor α. Hence any network activities α < β. Therefore, the transitory behaviour of the nodes is represented by *Tr*.

Whenever the forwarding behaviour of the node decreases, the Tr value will increase and otherwise the value will be decreased. The following algorithm represents deviation from Tr.

```
If $(PFB_{Aj}^{Ai}(T-1)) > PFB_{Aj}^{Ai}(T)$ then
        $Tr = Tr\text{-}1 + \alpha * (PFB_{Aj}^{Ai}(T-1) - PFB_{Aj}^{Ai}(T))$
If $(PFB_{Aj}^{Ai}(T-1)) < PFB_{Aj}^{Ai}(T)$ then
        $Tr = Tr\text{-}1 + \beta * (PFB_{Aj}^{Ai}(T-1) - PFB_{Aj}^{Ai}(T))$
        else
$Tr = Tr\text{-}1$
```

Finally, packet forwarding behaviours will be calculated based on the equation 3,

$$PFB_{Aj}^{Ai}(T) = PFB_{Aj}^{Ai}(T) * Tr \qquad (3)$$

The another level of trust used in the proposed method is called confidence level is represents the number of interactions between the trustor and the trustee. The following equations are used to indicate the confidence level of two nodes. The interactions can be measured by acknowledgement.

```
If (No.of interactions, high) then
        Confidence level is high
Otherwise
        Confidence level is low
```

The equation representation of the confident level trust can be calculated based on the following equation 4.

$$CL_{AiAj}(T) = \begin{cases} if\ (No.of\ iterations \geq Threshold\ value) \\ Confidence\ level\ is\ high, assume\ CL = 1 \\ Else\ if\ (No.of\ iteration = Threshold\ value) \\ Confidence\ level\ is\ moderate, assume\ CL = 0.5 \\ Else\ \ Confidence\ level\ is\ low, assume\ CL = 0.3 \end{cases} \qquad (4)$$

Then, Direct trust will be calculated by combining packet forwarding behaviour and confidence level. The following equation 5 is used to calculate the direct trust value.

$$DT_{Aj}^{Ai}(T) = \mu_1 PFB_{Aj}^{Ai}(T) + \mu_2 CL_{AiAj}(T) \qquad (5)$$

Where,

$DT_{Aj}^{Ai}(T)$ denotes the direct trust value of node Ai with respect to node Aj.

µ denoted the weighting factor and $\mu_1 + \mu_2 = 1$

E. Recommendation Trust

Sometimes the direct trust values will not be enough to assess the trustworthiness of the participating nodes. Hence, indirect trust i.e. recommendations from other nodes will also be considered. Direct trust values of a particular node may change over a period of time because of the significant features of IoT nodes. In that case, recommendation trust will be useful. Besides, a node may not have direct experience with other nodes and moreover an adversary node may act like a genuine node for one node and perform malicious activities for all other nodes in the network. Because of this reasons, a recommendation trust will be calculated as each node must have an interaction with other nodes during network operations. The following equation 6 is used to calculate recommendation trust.

$$RT_{Aj}^{Ai}(T)=\sum_{i=1}^{n}(DT_{Aj}^{Ai} * DT_{Am}^{Ai})/n \qquad (6)$$

In the above equation,

Where, i,j,m=1,2,3..n, i≠j and i≠m

$RT_{Aj}^{Ai}(T)$ denotes the recommendation trust of node Ai with respect to node Aj.

DT_{Aj}^{Ai} denotes the direct trust of node Ai with respect to node Am.

5. Proposed Algorithm

The following algorithm depicts the working principal of proposed algorithm

```
Algorithm
Input: Direct and Recommendation Trusts
Output: Classify nodes into Blackhole nodes, Trusted Nodes and Partially trusted nodes
  Begin:
If (performance of the network is well) then
        Continue the network operations
else
        for each (nodes evaluates every other nodes)
                Read Packet forwarding behaviour and
                Confidence level
                Calculate: Direct Trust (DT)
                Read direct trust of own and recommendation
                trust from other nodes
                Calculate: Recommendation Trust (RT)
Update in the trust table.
Read the Direct trust value and Recommendation Trust values of participating nodes in the IoT environment and make them as data points and set the behaviour status as Null for all the nodes.
Randomly select cluster heads or centroids from the overall data points.
Select the $DT_{Aj}^{Ai}(T)$ and $RT_{Aj}^{Ai}(T)$ of each node as one data point and randomly selected centroids as another data points.
Calculate the Euclidean distance between all the nodes and all centroid.
Assign the node to the closet centroid.
Repeat the process until all the nodes assign to the closest centroid.
Then form the cluster based on the centroids.
Repeat the process until newly formed cluster's centroid remains the same.
Stop the process.
Classify the nodes based on the cluster allocation into Blackhole nodes, trusted nodes and Partially trusted nodes.
Update status in the trust table.
Stop the process.
End
```

A. Mathematial Example

The following figure consists of eight nodes namely N1, N2, N3, N4, N5, N6, N7, N8 and N9 and these networks form an IoT environment. Each node consists of direct and indirect trust values of their neighbouring nodes. All these nodes are in the same communication range.

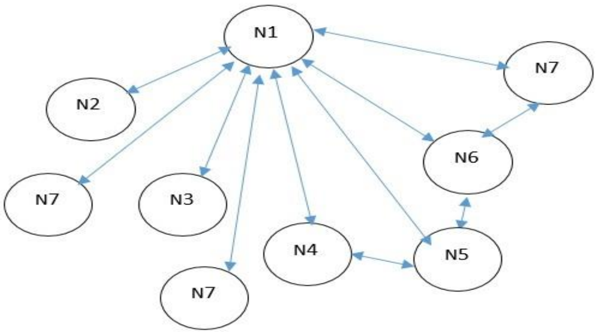

Figure 4 - Example Network

Assume node N1 is the evaluating node. It will assess the trustworthiness of all other nodes in the network. The trust table for node N1 is as follows. Initially, the behavioural status is null. After executing the KmeansT algorithm, node N1 can fill in the status. Hence, the initial status of the blacklist column will be set to NULL. Every table will also maintain their own trust values in its table.

TABLE II

Nodes' ID	DT	RT	Behaviour
N1	0.1	0.3	Null
N2	0.2	0.2	Null
N3	0.5	0.8	Null
N4	0.8	0.5	Null
N5	0.3	0.9	Null
N6	1	0.7	Null
N7	0.3	0.3	Null
N8	0.9	0.4	Null
N9	0.3	0.7	Null

b. Trust table of node N1 with respect to all other nodes

Step 1:

Initially the entire network is classified into two categories such as cluster 1, cluster 2 and cluster 3 as per the algorithm. Afterwards, we select three nodes as centroids or initial cluster heads for these clusters. Therefore, assume nodes N7, N9, and N8 are considered as centroids or cluster heads for cluster 1, cluster 2 and cluster 3 respectively. Table 3 represents the selection of centroids.

TABLE III

Cluster	DT	IT	Centroid
N7	0.3	0.3	(0.3,0.3)
N9	0.3	0.7	(0.3,0.7)
N8	0.9	0.4	(0.9,0.4)

c. Selection of centroid randomly

Step 2:

Calculate the Euclidean distance between the nodes and the centroids as per the algorithm. The below equation is used to calculate the Euclidean distance.

$$Euclidean\ Distance = \sqrt{(DT_x - DT_i)^2 - \left(RT_y - RT_i\right)^2}$$

In the above equation, DTx and RTy represents the direct trusts and recommendation trusts nodes in the network respectively and DTi and ITi are the randomly selected centroids.

Step 3:

Calculate the distance between the data points (Direct and Recommendation Trust) and Cluster heads (C1=, C2, C3).

Therefore, C1= (0.3, 0.3) N1= (0.1, 0.3)

C1N1=> $\sqrt{(0.1-0.3)^2+(0.3-0.3)^2}$ = 0.2

C2= (0.3, 0.7) N1= (0.1, 0.3)

C2N1=> $\sqrt{(0.1-0.3)^2+(0.3-0.7)^2}$ =0.4

C3= (0.9, 0.4) N1= (0.1, 0.3)

C3N1=> $\sqrt{(0.1-0.9)^2+(0.3-0.4)^2}$ =0.8

Similarly, for all other nodes in the network calculate the distance

TABLE IV

Nodes	C1 (0.3,0.3)	C2(0.3,0.7)	C3(0.9,0.4)	Cluster
N1 0.1 0.3	0.2	0.447214	0.806226	C1
N2 0.2 0.2	0.141421	0.509902	0.728011	C1
N3 0.5 0.8	0.538516	0.223607	0.565685	C2
N4 0.8 0.5	0.538516	0.538516	0.141421	C3
N5 0.3 0.9	0.6	0.2	0.781025	C2
N6 1 0.7	0.806226	0.7	0.316228	C3
N7 0.3 0.3	0	0.4	0.608276	C1
N8 0.9 0.4	0.608276	0.67082	0	C3
N9 0.3 0.7	0.4	0	0.67082	C2

d. Classification of clusters (Iteration 1)

Then,

Classify the clusters based on their names. Hence,
Cluster 1 => N1 (0.1, 0.3), N2 (0.2, 0.2), N7 (0.3, 0.3)
Cluster 2 => N3 (0.5,0.8), N5 (0.3, 0.9), N9 (0.3, 0.7)
Cluster 3=> N4 (0.8, 0.5), N6 (1.0.7), N8 (0.9,0.4)

Step 4:

Calculate the new centroids or cluster head by taking the mean of all the calculated data points from each cluster. Therefore,

The new cluster head of cluster 1 after iteration

1=> (0.1+0.2+0.3)/3, (0.3+0.2+0.3)/3 => (0.2, 0.266)

The new cluster head of cluster 2 after iteration

1=> (0.5+0.3+0.3)/3, (0.8+0.9+0.7)/3 => (0.366667, 0.8)

The new cluster head of cluster 3 after iteration

1=> (0.8+1+0.9)/3, (0.5+0.7+0.4)/3 => (0.9, 0.53333)

Now the iteration 1 is over. Then repeat the process with the new cluster heads.

TABLE V

Cluster	DT	IT	New Cluster Head or Centroids
C1	0.2	0.266	(0.2,0.266)
C2	0.366667	0.8	(0. 366667,0.8)
C3	0.9	0.53333	(0.9,0. 53333)

e. New cluster head information

TABLE VI

Nodes	C1 (0.2,0.266)	C2(0.3,0.7)	C3(0.9,0.4)	Cluster
N1 0.1 0.3	0.105622	0.566667	0.833332	C1
N2 0.2 0.2	0.066	0.622718	0.775312	C1
N3 0.5 0.8	0.6125	0.133333	0.480742	C2
N4 0.8 0.5	0.644016	0.527046	0.105408	C3
N5 0.3 0.9	0.641838	0.120185	0.703169	C2
N6 1 0.7	0.910141	0.641179	0.194368	C3
N7 0.3 0.3	0.105622	0.504425	0.643772	C1
N8 0.9 0.4	0.71271	0.666666	0.13333	C3
N9 0.3 0.7	0.445372	0.120185	0.622719	C2

f. Classification of clusters (Iteration 2)

Then,
Classify the clusters based on their names. Hence,

Cluster 1 => N1 (0.1, 0.3), N2 (0.2, 0.2), N7 (0.3, 0.3)
Cluster 2 => N3 (0.5,0.8), N5 (0.3, 0.9), N9 (0.3, 0.7)
Cluster 3=> N4 (0.8, 0.5), N6 (1.0.7), N8 (0.9,0.4)

Calculate the new centroids or cluster head by taking the mean of all the calculated data points from each cluster.

Therefore,
The new cluster head of cluster 1 after iteration 2=> (0.1+0.2+0.3)/3, (0.3+0.2+0.3)/3 => (0.2, 0.266)

The new cluster head of cluster 2 after iteration 2=> (0.5+0.3+0.3)/3, (0.8+0.9+0.7)/3 => (0.366667, 0.8)

The new cluster head of cluster 3 after iteration 2=> (0.8+1+0.9)/3, (0.5+0.7+0.4)/3 => (0.9, 0.53333)

The algorithm will stop as two iterations got the same cluster heads or centroids therefore no more iterations. Hence, the classified clusters will be considered as final. Based on the algorithm, nodes can be classified into three categories such as Blackhole nodes, Trusted nodes and Partially trusted nodes.

From the below table, nodes N1, N2 and N8 will be considered as blackhole nodes and those nodes will be eliminated from the network. Only trusted and partially trusted nodes will be allowed to participate in network activations.

TABLE VII

Nodes	DT	RT	Iteration 1	Iteration 2	Behaviour
N1	0.1	0.3	C1	C1	Blackhole
N2	0.2	0.2	C1	C1	Blackhole
N3	0.5	0.8	C2	C2	Partially Trusted
N4	0.8	0.5	C3	C3	Trusted
N5	0.3	0.9	C2	C2	Partially Trusted
N6	1	0.7	C3	C3	Trusted
N7	0.3	0.3	C1	C1	Blackhole
N8	0.9	0.4	C3	C3	Trusted
N9	0.3	0.7	C2	C2	Partially Trusted

g. Nodes behavioral Status

6. RESULTS AND DISCUSSION

The proposed KmeansT has compared with traditional RPL routing protocol and Trust-based RPL in terms of various performance metrics like packet delivery ratio, detection ratio and end to end delay. The following table depicts the simulation parameter settings:

TABLE VIII

Simulation Parameters	
Simulation Tool	InstantContiki/Cooja 3.0
Total simulation runtime	1800 Seconds
Area covered by the simulation	100m X 100m
Mote Type	Tmote Sky
Range of Interferences	100m
No.of nodes	70 (Max)
Sink (Root Node)	1
Blackhole nodes	5-20
Legitimate nodes	>20 and <=70
Deployment Environment	General
Network Protocol	IP
Routing Protocol	RPL
Wireless Transmission Range	50 meters
Traffic Rate	1 packet sent every 10 sec
Radio Medium model	UDGM Distance Loss

h. Simulation Parameters

A. *Detection Ratio*

This parameter determines the ability of an algorithm to detect adversaries i.e. blackhole nodes. The traditional RPL routing protocol does not have any detection ability in the design of the protocol itself therefore it does not take into account. While the trust-based RPL protocol makes use of a single metric to evaluate the trust worthiness of nodes therefore the detection ratio is decreased due to its weak measurement when blackhole nodes are increased. At the same time, KmeansT makes use of both direct trust and Recommendation trust to evaluate the trustworthiness besides K means clustering algorithm effectively identify the blackhole nodes therefore though the number of blackhole nodes are increased the detection of blackhole nodes are constantly increasing compare with Trust-based RPL. The average detection ratio of Trust-based RPL is 19.54% where as KmeansT gains 35.38% when blackhole nodes are randomly placed. During the simulation the detection ratio of KmeansT reached up to 89.42%. The following "Fig. 5" depicts the detection ratio analysis.

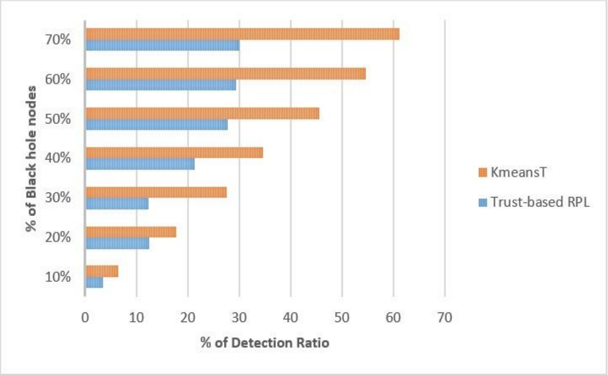

Figure 5 - Detection ratio versus Blackhole nodes

B. Packet Delivery Ratio

The impact of packet delivery ratio has been analysed in the “Fig. 6”.

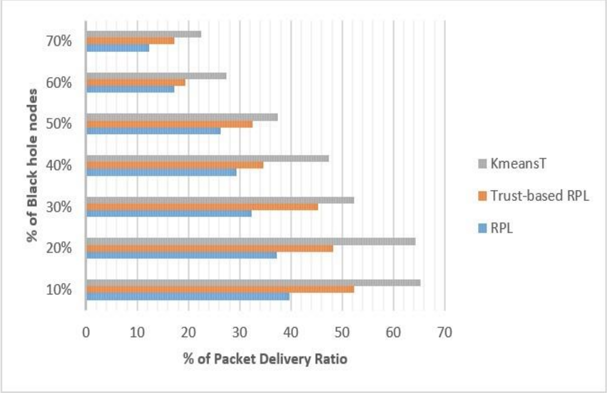

Figure 6 - Packet Delivery Ratio versus Blackhole nodes

This factor determines the number of packets being received by the destination nodes over the actual number of packets that are being transmitted by the sender node. From the simulation results we observed that the packet delivery ratio reached around 97.5% when there are no adversaries. But when black hole nodes are added at a regular interval of time, the packet delivery ratio is significantly decreased. The packet delivery ratio of KmeansT is high compare with other two protocols as the K-means algorithm plays a major role in detecting blackhole nodes. When the next iteration starts, those nodes can be eliminated from the network therefore higher packet delivery ratio compare with others. Therefore, KmeansT gives an 82.5% delivery ratio even with 50% of malicious nodes in the network. As a single metric used in Trust-based RPL, packet delivery ratio is low compared with KmeansT and high compared with RPL. The traditional RPL does not have the ability to detect the black hole nodes therefore packet delivery ratio is low in the presence of black hole nodes.

C. End to end delay

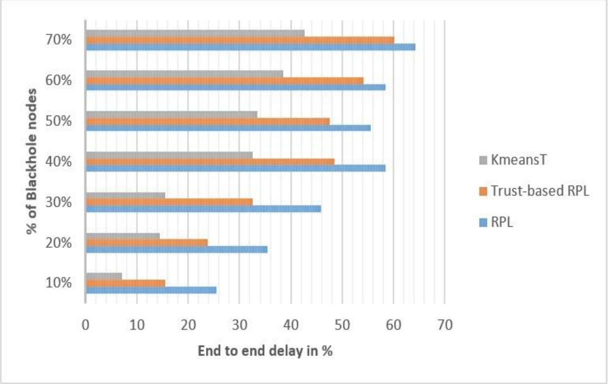

Fig. 7 - End to End delay versus Blackhole nodes

The below “Fig.7” represents the end-to-end delay of analysis. This metric deals with the time taken for a packet to reach the destination node from the source node. The end to end delay is low for KmeansT model compare with others two as blackhole nodes are eliminated with the help of multiple trust evaluation mechanism therefore the presence of blackhole nodes is low therefore lower end to end delay. While in the trust-based RPL single metric is used therefore the presence of blackhole nodes is possible therefore higher end to end delay compare with KmeansT model at the same time lower than RPL. The RPL has higher end-to-end delay as the presence of blackholes is high. The average end to end delay is 26.28% for KmeansT whereas for RPL it is 49.07% and Trust-based RPL is 35.2%.

7. CONCLUSION

Security is a main concern in the IoT environment as it deals with sensitive information. The IoT environment is affected by various attacks. Blackhole attack is a kind of attack to drop all the incoming packets by the way the performance of the entire network becomes degraded. To address this, many security mechanisms have been proposed however they might not be suitable for a resource constrained IoT environment. *Moreover*, there is no complete detection of such attacks during network activities. To address this, the K-meansT algorithm is proposed in this work. The main advantage of this method is to detect black hole attacks accurately with the help of trust management along with a K-means clustering algorithm.

Since this algorithm did not use complex algorithms, it is quite suitable for a resource consuming IoT environment. Moreover, this algorithm deals with core components of routing such as packet forwarding behaviour as one of the key metrics when evaluating the trustworthiness of nodes. In addition, k means clustering algorithm classifies the nodes as trusted, partially trusted and blackhole nodes. Except blackhole nodes, all other nodes will participate in the network activities therefore the performance of the network will be maintained. In the future, the work will be focused on some other attacks which are affecting the IoT environment.

REFERENCES

[1] P. K. Sadhu, V. P. Yanambaka, and A. Abdelgawad, “Internet of Things: Security and Solutions Survey,” Sensors, vol. 22, no. 19, p. 7433, Sep. 2022, doi: 10.3390/s22197433.

[2] B. Kaur et al., “Internet of Things (IoT) security dataset evolution: Challenges and future directions,” Internet Things, vol. 22, p. 100780, Jul. 2023, doi: 10.1016/j.iot.2023.100780.

[3] M. Litoussi, N. Kannouf, K. El Makkaoui, A. Ezzati, and M. Fartitchou, “IoT security: challenges and countermeasures,” Procedia Comput. Sci., vol. 177, pp. 503–508, 2020, doi: 10.1016/j.procs.2020.10.069.

[4] I. Ahmad, M. S. Niazy, R. A. Ziar, and S. Khan, “Survey on IoT: Security Threats and Applications,” J. Robot. Control JRC, vol. 2, no. 1, 2021, doi: 10.18196/jrc.2150.

[5] M. Liyanage, A. Braeken, P. Kumar, and M. Ylianttila, IoT Security: Advances in Authentication. John Wiley & Sons, 2020.

[6] A. O. Affia, A. Nolte, and R. Matulevičius, “IoT Security Risk Management: A Framework and Teaching Approach,” Inform. Educ., Apr. 2023, doi: 10.15388/infedu.2023.30.

[7] S. Rekha, L. Thirupathi, S. Renikunta, and R. Gangula, “Study of security issues and solutions in Internet of Things (IoT),” Mater. Today Proc., vol. 80, pp. 3554–3559, 2023, doi: 10.1016/j.matpr.2021.07.295.

[8] R. Ahmad and I. Alsmadi, “Machine learning approaches to IoT security: A systematic literature review,” Internet Things, vol. 14, p. 100365, Jun. 2021, doi: 10.1016/j.iot.2021.100365.

[9] E. Schiller, A. Aidoo, J. Fuhrer, J. Stahl, M. Ziörjen, and B. Stiller, “Landscape of IoT security,” Comput. Sci. Rev., vol. 44, p. 100467, May 2022, doi: 10.1016/j.cosrev.2022.100467.

[10] I. H. Sarker, A. I. Khan, Y. B. Abushark, and F. Alsolami, “Internet of Things (IoT) Security Intelligence: A Comprehensive Overview, Machine Learning Solutions and Research Directions,” MATHEMATICS & COMPUTER SCIENCE, preprint, Mar. 2022. doi: 10.20944/preprints202203.0087.v1.

[11] H. P. Ng, S. H. Ong, K. W. C. Foong, P. S. Goh, and W. L. Nowinski, “Medical Image Segmentation Using K-Means Clustering and Improved Watershed Algorithm,” in 2006 IEEE Southwest Symposium on Image Analysis and Interpretation, Denver, CO: IEEE, 2006, pp. 61–65. doi: 10.1109/SSIAI.2006.1633722.

[12] M. N. Reza, I. S. Na, S. W. Baek, and K.-H. Lee, “Rice yield estimation based on K-means clustering with graph-cut segmentation using low-altitude UAV images,” Biosyst. Eng., vol. 177, pp. 109–121, Jan. 2019, doi: 10.1016/j.biosystemseng.2018.09.014.

[13] N. Nidheesh, K. A. Abdul Nazeer, and P. M. Ameer, “An enhanced deterministic K-Means clustering algorithm for cancer subtype prediction from gene expression data,” Comput. Biol. Med., vol. 91, pp. 213–221, Dec. 2017, doi: 10.1016/j.compbiomed.2017.10.014.

[14] H. Cui, G. Ruan, J. Xue, R. Xie, L. Wang, and X. Feng, “A collaborative divide-and-conquer K-means clustering algorithm for processing large data,” in Proceedings of the 11th ACM Conference on Computing Frontiers, Cagliari Italy: ACM, May 2014, pp. 1–10. doi: 10.1145/2597917.2597918.

[15] Md. Jahiruzzaman and A. B. M. Aowlad Hossain, “Detection and classification of diabetic retinopathy using K-means clustering and fuzzy logic,” in 2015 18th International Conference on Computer and Information Technology (ICCIT), Dhaka: IEEE, Dec. 2015, pp. 534–538. doi: 10.1109/ICCITechn.2015.7488129.

[16] A. Musaddiq, Y. B. Zikria, Zulqarnain, and S. W. Kim, “Routing protocol for Low-Power and Lossy Networks for heterogeneous traffic network,” EURASIP J. Wirel. Commun. Netw., vol. 2020, no. 1, p. 21, Dec. 2020, doi: 10.1186/s13638-020-1645-4.

[17] M. Conti, P. Kaliyar, M. M. Rabbani, and S. Ranise, “SPLIT: A Secure and Scalable RPL routing protocol for Internet of Things,” in 2018 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Limassol: IEEE, Oct. 2018, pp. 1–8. doi: 10.1109/WiMOB.2018.8589115.

[18] T. Kothmayr, C. Schmitt, W. Hu, M. Brünig, and G. Carle, “DTLS based security and two-way authentication for the Internet of Things,” Ad Hoc Netw., vol. 11, no. 8, pp. 2710–2723, Nov. 2013, doi: 10.1016/j.adhoc.2013.05.003.

[19] P. N. Mahalle, N. R. Prasad, and R. Prasad, “Threshold Cryptography-based Group Authentication (TCGA) scheme for the Internet of Things (IoT),” in 2014 4th International Conference on Wireless Communications, Vehicular Technology, Information Theory and Aerospace & Electronic Systems (VITAE), Aalborg, Denmark: IEEE, May 2014, pp. 1–5. doi: 10.1109/VITAE.2014.6934425.

[20] R. Sahay, G. Geethakumari, B. Mitra, and V. Thejas, “Exponential Smoothing based Approach for Detection of Blackhole Attacks in IoT,” in 2018 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Indore, India: IEEE, Dec. 2018, pp. 1–6. doi: 10.1109/ANTS.2018.8710073.

[21] A. A. Alamr, F. Kausar, J. Kim, and C. Seo, “A secure ECC-based RFID mutual authentication protocol for internet of things,” J. Supercomput., vol. 74, no. 9, pp. 4281–4294, Sep. 2018, doi: 10.1007/s11227-016-1861-1.

[22] D. Airehrour, J. A. Gutierrez, and S. K. Ray, “A Trust-Aware RPL Routing Protocol to Detect Blackhole and Selective Forwarding Attacks,” J. Telecommun. Digit. Econ., vol. 5, no. 1, pp. 50–69, Mar. 2017, doi: 10.18080/jtde.v5n1.88.

[23] Z. A. Khan and P. Herrmann, “A Trust Based Distributed Intrusion Detection Mechanism for Internet of Things,” in 2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA), Taipei, Taiwan: IEEE, Mar. 2017, pp. 1169–1176. doi: 10.1109/AINA.2017.161.

[24] V. Sharma, I. You, K. Andersson, F. Palmieri, M. H. Rehmani, and J. Lim, “Security, Privacy and Trust for Smart Mobile-Internet of Things (M-IoT): A Survey,” IEEE Access, vol. 8, pp. 167123–167163, 2020, doi: 10.1109/ACCESS.2020.3022661.

[25] H. Perrey, M. Landsmann, O. Ugus, T. C. Schmidt, and M. Wählisch, “TRAIL: Topology Authentication in RPL.” arXiv, Dec. 15, 2015. Accessed: Nov. 01, 2023. [Online]. Available: http://arxiv.org/abs/1312.0984

[26] T. Zhang, T. Zhang, X. Ji, and W. Xu, “Cuckoo-RPL: Cuckoo Filter based RPL for Defending AMI Network from Blackhole Attacks,” in 2019 Chinese Control Conference (CCC), Guangzhou, China: IEEE, Jul. 2019, pp. 8920–8925. doi: 10.23919/ChiCC.2019.8866139.

[27] D. Airehrour, J. A. Gutierrez, and S. K. Ray, “SecTrust-RPL: A secure trust-aware RPL routing protocol for Internet of Things,” Future Gener. Comput. Syst., vol. 93, pp. 860–876, Apr. 2019, doi: 10.1016/j.future.2018.03.021.

[28] A. Tandon and P. Srivastava, “Trust-based Enhanced Secure Routing against Rank and Sybil Attacks in IoT,” in 2019 Twelfth International Conference on Contemporary Computing (IC3), Noida, India: IEEE, Aug. 2019, pp. 1–7. doi: 10.1109/IC3.2019.8844935.

[29] J. Kaur and G. Singh, “A Blockchain-Based Machine Learning Intrusion Detection System for Internet of Things,” in Principles and Practice of Blockchains, K. Daimi, I. Dionysiou, and N. El Madhoun, Eds., Cham: Springer International Publishing, 2023, pp. 119–134. doi: 10.1007/978-3-031-10507-4_6.

Mr. Shameer Mohammed received his MCA from the University of Madras, Chennai. He's pursuing a PhD in computer science in Internet of Things at Karpagam Academy of Higher Education in Coimbatore, Tamilnadu, India. He has extensive experience with academic institutions and IT firms. IoT, Big Data Analytics, Software engineering, web technologies and Microservices are his research interests. He gave workshops, developed and published research papers in reputable publications, and provided community service programmes.

Dr. L. Gnanaprasanambikai is currently working as Assistant Professor at Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India. Her research interests include network security, soft skills. She has published various articles in reputed journals.