K-Means Clustering-Based Trust (KmeansT) Evaluation Mechanism for Detecting Blackhole Attacks in IoT Environment

Shameer M.¹ and Gnanaprasanambikai L.²

¹, ²Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore, Tamilnadu, India

Received 13 Dec. 2023, Revised 25 Apr. 2024, Accepted 1 May 2024, Published 1 Aug. 2024

Abstract: The Internet of Things (IoT) has revolutionized numerous aspects of our lives, offering many applications that enhance convenience and comfort. However, alongside its significant benefits, IoT introduces several research challenges, with security emerging as a primary concern. Given the sensitive nature of the information exchanged within IoT environments, ensuring robust security measures is imperative. One prominent threat in IoT environments is the potential for malicious attacks, which can exploit vulnerabilities and disrupt network operations. Among these threats, blackhole attacks pose a particularly concerning risk, as they involve malicious entities dropping all incoming packets, disrupting routing operations, and impeding communication. To mitigate the risks posed by blackhole attacks and enhance the security of IoT networks, a novel approach known as the K-means clustering-based Trust (KmeansT) evaluation mechanism has been proposed. This innovative method employs a multifaceted trust evaluation process, incorporating both direct observations and recommendations from other network entities. By leveraging the K-means clustering algorithm, the proposed mechanism enhances the effectiveness of trust evaluation, enabling a more accurate assessment of node reliability and integrity. One of the key strengths of the KmeansT approach lies in its ability to identify and mitigate blackhole attacks within the IoT environment effectively. Through rigorous mathematical modeling and simulation studies, the efficacy of the proposed mechanism in detecting and neutralizing blackhole threats is demonstrated. Simulation results are analyzed comprehensively, with performance metrics compared against existing models to assess the effectiveness of the KmeansT approach. By evaluating constraints such as end-to-end delay, packet delivery, and detection ratio, the superiority of the anticipated mechanism in safeguarding IoT networks against blackhole attacks is underscored.

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. Consequently, information and communication technology plays a major role in satisfying consumer needs. The Internet of Things is one such protruding 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 preceding two extents have experienced a sturdy rise in the fabrication and deployment of sensing-and connectivity-enabled electronic devices, swapping “regular” physical objects. The ensuing Internet of Things (IoT) will soon become obligatory for many application domains. Smart objects are unremittingly cohesive within factories, cities, buildings, health institutions, and private homes. Nearly 30 years after the birth of IoT, society is confronted with striking trials regarding IoT security. Due to the interconnectivity and ubiquitous
use of IoT devices, cyberattacks have extensive impacts on multiple stakeholders. Ancient threats spectate that the IoT domain holds various susceptibilities, oppressed 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 from 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 interoperability problems that lead 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 1 depicts the security issues of IoT [6], [7], [8].

The security issues are broadly ordered into three major sorts: high-level, mid-level, and low-level. The extraordinary retreat issues occur in the application layer, mid-level security issues arise in both transport and network layers and low-level security issues occur 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 systems, blockchain technology, symmetric and asymmetric cryptography algorithms, hash functions, etc. are effective in providing and ensuring security in the network layer. However, they might not apply to resource-constrained IoT devices. Applying all those algorithms in resource-constrained IoT devices leads to security violations [6], [7], [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 neighbor discovery, duplication or replay attack, wormhole attack, blackhole attack, sinkhole attack, Sybil attack, and buffer reservation attacks affect the network layer commonly. The proposed research work focuses on black hole attacks. It is a kind of attack that affects 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. The result is the overall performance of the network becomes degraded. The network layer’s main concern is routing. In an IoT environment, data or control packets are 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 security the following security requirements must be considered such as confidentiality, authentication, authorization, access control, and non-repudiation among requirements authentication is considered as a primary requirement in the security aspects of the IoT environment as it safeguards the preliminary level of security. However, in the IoT environment ensuring authentication is an intricate task as IoT has diverse devices, cross-platform capabilities, and resource-constrained nature. Therefore, the IoT environment is expecting proper authentication along with a secure mechanism to protect the IoT environment [10].

A. Research Objectives and Contribution

Ensuring security within the realm of IoT presents a noteworthy challenge, garnering considerable attention from researchers. This augmented motivation stems from the inherent limitations in IoT devices’ processing, storage, and battery capabilities, which serve as vulnerabilities susceptible to various attacks. The diverse nature of these devices, in particular, exacerbates interoperability issues, thereby increasing the risk of security breaches. Among these threats, the black hole attack stands out as particularly concerning, habitually perpetrated by compromised internal nodes. We propose the K-means clustering-based trust (KmeansT) evaluation mechanism to address this challenge. This approach addresses the problem by leveraging trust management principles and the K-means clustering algorithm to meritoriously detect and neutralize black hole attacks.

Below is the outline for the remainder of the paper: Section 2 is about contextual information and includes the k-means clustering algorithm, the influence of a blackhole occurrence on the RPL routing protocol, and an explanation of how RPL works; Section 3 delves into SLR; Section 4 describes the proposed model; Section 5 shows the proposed algorithm; Section 6 is about outcomes and discussion; section 7 is the inference and the last section is limitations and future research directions.
2. Background

The proposed algorithm makes use of the K-means clustering algorithm to ascertain the black hole attack over the RPL routing protocol. The ensuing part discusses the k-means clustering algorithm.

A. K-Means Clustering Algorithm

It is the easiest and has less computational overhead [11]. This simple algorithm is used to categorize the data into K clusters. These clusters are depicted by the centroids [12]. The recognized K-means group is 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 have been designated as input with a collection of information points [13].

The procedure starts by starting K-cluster centers. These centers were chosen randomly or based on some heuristic procedure [14]. The center is called the prototype point (centroids). The data points after the data set have been allotted to individual clusters based on the closest prototype point. Then, the mean data points are calculated by taking the average of the data point’s coordinate values for separate collection. The mean arguments comprise a new set of prototype points. Again, every data point is allotted to a cluster of its closest prototype points. This phase of the group is conclusive clustering results. The Euclidean distance has been used for proximity measure in K-means. This algorithm consists of many advantages that variety it very familiar. The significant ones are simplicity and easy implementation. Because 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:

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

B. The impact of Blackhole over RPL

As discussed in the introduction section, the routing protocols could be used to route the information from one place to another place. In IoT many routing protocols have been used, however, Routing Protocol for Low Power Lossy Networks (RPL) is habitually employed in IoT environments. The suggested model incorporates the RPL routing protocol, as explained in detail in reference [16]. The black hole attack is one kind of attack that harms the routing operation by dropping all the incoming packets that are dedicated to forward to others. Figure 2 denotes a network structure with no blackhole attack and Figure 3 denotes a network structure with blackhole attack.

In an RPL-based IoT environment, initially, the devices are authenticated and trusted hence DODAG construction has been done without any difficulty. Over a while, the behavior of devices might be changed and perform malicious activities like black hole attacks. Figure 3 represents the typical RPL network with a black hole attack respectively. In Figure 3, an IoT environment consists of 10 nodes along with the origin node. Here, node/device 3 is assumed a black hole device. Hence, it publicizes the situation that is taking the direct pathway to reach the root device R. Therefore, device 7 assumes that device 3 has a path to the root node and forwards its packet to that node. As device 3
is a block hole node, it will not forward the packet to root node R, in addition, it reduces all the inward packages that are envisioned to onward. By the way, a black hole attack is executed in the IoBT network. The subsequent section will delve into a systematic literature review discussion.

3. REVIEW OF LITERATURE

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

In [17], the authors introduced SRAP, a Routing Protocol for Low-Power and Lossy Networks (RPL)-based solution tailored for non-homogeneous IoT environments. SRAP is designed to enhance productivity while addressing scalability concerns by minimizing overhead. One notable feature is its utilization of Destination Advertisement Objects (DAO) in an encrypted manner, effectively thwarting potential threats posed by malicious devices within the network.

In [18], the authors proposed a security model grounded in the Rivest Shamir Adleman (RSA) public key cryptography algorithm. This model is specifically engineered to fortify RPL-based routing protocols by ensuring confidentiality, integrity, and authentication. By leveraging RSA encryption, the model safeguards communication channels against unauthorized access and tampering, thereby upholding the essential security requirements of IoT networks.

In [19], the authors propose a model to combat blackhole attacks utilizing an exponential smoothing algorithm. This innovative approach aims to address the topological disruptions caused by blackhole nodes by calculating packet delivery times from the root node. Leveraging this data, the algorithm makes informed decisions to mitigate blackhole attacks, effectively safeguarding network integrity and ensuring uninterrupted communication flow within the IoT environment.

In [20], the authors proposed a robust security model grounded in Elliptic Curve Diffie-Hellman (ECDH) cryptography, which ensures multiple security properties essential for IoT environments. These properties include confidentiality, authentication, ambiguity resolution, location privacy preservation, and data packet forwarding security. By leveraging ECDH cryptography, the model provides a comprehensive framework for safeguarding IoT networks against various security threats while maintaining privacy and integrity.

In [21], the authors introduced a trust prototype based on fuzzy logic, aiming to combat blackhole attacks by establishing trusted routes within IoT networks. This model utilizes fuzzy logic to evaluate trust levels, facilitating the formation of reliable routes that mitigate the risks posed by malicious entities. By leveraging fuzzy logic, the model adapts to dynamic network conditions and effectively identifies trustworthy paths, thereby enhancing the security and reliability of IoT communication.

Addressing the challenges posed by gray hole and worm hole attacks, [22] presents a model that focuses on trust-based mechanisms built on forwarding and ranking checks. By evaluating the trustworthiness of nodes based on their behavior in packet forwarding, this model effectively identifies and mitigates the risks associated with these types of attacks, ensuring the integrity and reliability of IoT networks.

In [23], the authors proposed a trust model grounded in energy considerations, leveraging local trust design and parent node feedback to evaluate node trustworthiness. By assessing nodes based on their energy usage and collecting opinions from parent nodes, this model provides a reliable framework for establishing trust relationships within IoT networks, thereby enhancing security and reliability.

In [24], a lightweight cuckoo filter-based security mechanism is proposed to address blackhole attacks. This model employs secure rank calculation, infrastructure establishment, and node registration stages to evaluate node authenticity and permit only authenticated nodes to participate in network operations, thereby mitigating the risks associated with blackhole attacks.

Mitigating rank and Sybil attacks, [25] presents a trust model that evaluates the dependability and reliability of nodes. By incorporating positive acknowledgments and employing a fuzzy threshold mechanism to broadcast trust values across networks, the aforementioned model effectively pinpoints and quarantines malicious nodes, enhancing the security and reliability of IoT networks.

In [26], a model leveraging context awareness is proposed to address Sybil and rank assaults in RPL-based IoT networks. By assessing the reliability of both child and parent nodes and computing direct and indirect trust values based on various parameters, including energy, rank value, hop count, and node behavior, this model establishes trustworthy communication paths, thereby enhancing network security.

In [27], resource-constrained lightweight symmetric ciphers are used. Popular Arduino and Raspberry Pi were tested. They cast off an ATMEGA328p microcontroller to ripen 39 block ciphers for different block and key sizes and tested their encryption and decryption performance, cost, and energy efficacy also added 80 second-round NIST stream and block cipher algorithms to the previously studied
ciphers. This extensive study scrutinized dormancy and energy adeptness for all ciphers with alike block and key sizes.

In [28], the authors compare city and highway IoV network speeds using random waypoint and Manhattan grid network mobility models for static and mobile nodes. These findings provide light on dynamic IoV networks’ stability, scalability, and reliability and suggest strategies to improve hybrid objective functions with Quality of Service for scalable networks.

In [29], the author covered the DIS flooding attack debated in this paper portentous that swelling the numeral of attacker nodes has an evident negative effect on end-to-end delay, packet detection ratio, and power consumption.

Moving forward, the authors in [30], scrutinize routing attacks and suggest mitigation in RPL-based IoT networks. It presents a new classification framework that links RPL assaults to their response methods, improving security strategy knowledge and organization.

Authors swell cases with copious sinks and attackers in [31] to improve network attack detection. Overall, the research sheds clarity on how different assaults affect RPL setups and proposes network security improvements.

In [32], Lightweight Cryptography (LWC) is endorsed for resource-limited IoT devices. Lightweight cryptography rallies IoT network security. Security liabilities in IoT systems are premeditated. It observes lightweight block, stream, hash, and Elliptic Curve Cryptography.

Authors in [33], booms a safe visual cryptography-based mutual authentication scheme. The etiquette uses visual cryptography to encrypt and decrypt secret images and tickets for mutual authentication. Accessing cloud amenities entails a ticket from the authentication server. Authentication practices three mutual secret keys for encryption and decryption.

Authors in [34], offered an IoT archetype, EGCrypto, that is reliable and efficient. Elliptic Galois cryptography and matrix XOR steganography protect EGCrypto. Self-adaptive differential evolution, fitness and diversity ranking, zonal control-specific development, and adaptive mutation enhance performance. Optimizing EGCrypto hyperparameters improves efficiency and efficacy. EGCrypto’s elliptic Galois cryptography safeguards IoT data. In picture cover blocks, optimization furts encrypted data. This technique decrypts and recovers IoT data at the receiving end, ensuring safe transfer.

Finally, [35] suggests a security mechanism based on blockchain technology and machine learning algorithms for intrusion detection, aiming to eliminate internal attacks within IoT environments. By leveraging the immutability and transparency of blockchain and the analytical capabilities of machine learning, this model offers a robust framework for detecting and mitigating internal security threats, thereby enhancing the overall security posture of IoT networks.

A. Research Gap:
Despite the success of the security methods presented in previous discussions, there remains a need for more effective mechanisms due to their limitations. Most of the security techniques proposed are cryptographic, and they may be unsuitable for Internet of Things devices due to the resource limitation constraint. In addition, they mostly concentrate on generic security concerns rather than focusing on specific attack types. The prevalent routing safekeeping methods and the innovative trust organization proposed in this research illustrate the importance of both the fundamentals and the likelihood of this proposed work in addressing the above-described concerns. The next component of this research illuminates the proposed model.

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

A. Assumptions:
• The network environment comprises 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 partaking devices in the environment are resource-constrained in the amount of energy, memory, and processing capabilities.

• The network entails fewer amount 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 they try to save their energy.

• Every node maintains a trust table where all the confidence-related evidence of the participating nodes can be stored.

• The proposed model will execute over some time or when the performance of the model decreases.

The following Table I illustrates the schema layout of a table along with sample data.

<table>
<thead>
<tr>
<th>Node’s ID</th>
<th>DT</th>
<th>RT</th>
<th>Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>0.1</td>
<td>0.3</td>
<td>partially trusted</td>
</tr>
</tbody>
</table>
Nodes’ ID - Nodes Identification
DT – Direct Trust
RT- Recommendation Trust
Behavior - Blackhole node or Genuine Node

B. Direct Trust
Over some 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 \( A_i \) wants to estimate the direct trust of node \( A_j \) with the help of the following two factors such as packet forwarding behavior and confidence level. The packet forwarding behavior 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, it will not forward the packets as it tries to save their energy. Within this consideration, the following equation is used to calculate packet forwarding behavior.

\[
PFB^{A_i}_{A_j}(T) = \frac{P\_F(T)_{A_i}}{P\_D(T)_{A_i} + P\_F(T)_{A_i}}
\]

Where,

- \( PFB^{A_i}_{A_j}(T) \) denotes the packet forwarding behavior of node \( A_i \) concerning node \( A_j \).
- \( P\_F(T)_{A_i} \) denoted the packet forwarding ratio of node \( A_i \).
- \( P\_D(T)_{A_i} \) denotes the packet dropping ratio of node \( A_i \).
- \( T \) represents the time.

During the transmission, the blackhole nature of IoT devices makes them selfish. In that case, those nodes will not perform or be involved in any network operations. This is important to analyze the behavior of the nodes over time. If the mobile nodes perform well that can be represented by the praise factor \( \beta \) otherwise it is not performing well can be represented by the penalty factor \( \alpha \). Hence any network activities \( \alpha < \beta \). Therefore, the transitory behavior of the nodes is represented by \( Tr \).

Whenever the forwarding behavior of the node decreases, the \( Tr \) value will increase and otherwise the value will be decreased. The following algorithm represents a deviation from \( Tr \).

If \( PFB^{A_i}_{A_j}(T - 1) > PFB^{A_i}_{A_j}(T) \) then
\[
Tr = Tr - 1 + \alpha * (PFB^{A_i}_{A_j}(T - 1) - PFB^{A_i}_{A_j}(T))
\]
If \( PFB^{A_i}_{A_j}(T - 1) < PFB^{A_i}_{A_j}(T) \) then
\[
Tr = Tr - 1 + \beta * (PFB^{A_i}_{A_j}(T - 1) - PFB^{A_i}_{A_j}(T))
\]
else \( Tr = Tr - 1 \)

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

\[
PFB^{A_i}_{A_j}(T) = PFB^{A_i}_{A_j}(T) * Tr
\]

The other 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 acknowledgment.

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.

\[
CL_{A_i,A_j}(T) = \begin{cases} 
\text{If (No. of iterations } \geq \text{ Threshold value) then} \\
\text{Confidence level is high, assume } CL=1 \\
\text{else If (No. of iterations } = \text{ Threshold value) then} \\
\text{Confidence level is moderate, assume } CL=0.5 \\
\text{else Confidence level is low, assume } CL=0.3 
\end{cases}
\]

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

\[
DT^{A_i}_{A_j}(T) = \mu_1 PFB^{A_i}_{A_j}(T) + \mu_2 CL_{A_i,A_j}(T)
\]

Where,

- \( DT^{A_i}_{A_j}(T) \) denotes the direct trust value of node \( A_i \) with respect to node \( A_j \).
- \( \mu_1 \) denoted the weighting factor and \( \mu_1+\mu_2 = 1 \)

E. Recommendation Trust
Sometimes the direct trust values will not be ample to assess the fidelity of the participating nodes. Hence, secondary trust i.e. recommendations from other nodes will also be considered. Direct trust values of a particular node may change over 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 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 these 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.

\[ RT_{A_iA_j}^{\text{(T)}} = \frac{\sum_{i=1}^{n} (DT_{A_iA_j}^{\text{(T)}} * DT_{A_iA_m}^{\text{(T)}})}{n} \]  

where, \( i, j, m = 1, 2, 3, \ldots, n \), \( i \neq j \) and \( i \neq m \)

\( RT_{A_iA_j}^{\text{(T)}} \) denotes the recommendation trust of node \( A_i \) regarding node \( A_j \).

\( DT_{A_iA_j}^{\text{(T)}} \) signifies the direct confidence of node \( A_i \) concerning node \( A_m \).

The following section unfolds the mathematical and simulated approaches and methodology used in this study.

5. Proposed Algorithm

The following depicts the proposed algorithm’s working principle:

A. Research Methodology

This experimental methodology demarcates the sequential procedures for assessing network performance, executing trust evaluation, establishing clusters based on trust metrics, categorizing nodes, and revising trust status. The objective of the method is to determine and categorize nodes according to their reliability to reduce the potential vulnerabilities caused by blackhole attacks in IoT environments.

B. Mathematical Example

The following Figure 4. Example Network 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 neighboring nodes. All these nodes are in the same communication range.

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 behavioral 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 its trust values in its table. Table II indicates the trust table of node N1 concerning all other nodes in the network.

Step 1: Initially, the entire network is classified into two categories such as cluster 1, cluster 2, and cluster 3 as per the algorithm. Afterward, we select three nodes as centroids or initial cluster heads for these clusters. Therefore, assume

![Figure 4. Example Network](https://journal.uob.edu.bh/)
nodes N7, N9, and N8 are considered centroids or cluster heads for cluster 1, cluster 2, and cluster 3 respectively. Table III represents the selection of centroids.

**Table III. Selection of Centroid Randomly**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>DT</th>
<th>IT</th>
<th>Centroid</th>
</tr>
</thead>
<tbody>
<tr>
<td>N7</td>
<td>0.3</td>
<td>0.3</td>
<td>(0.3,0.3)</td>
</tr>
<tr>
<td>N9</td>
<td>0.3</td>
<td>0.7</td>
<td>(0.9,0.4)</td>
</tr>
</tbody>
</table>

**Step2:** Calculate the Euclidean distance between the nodes and the centroids as per the algorithm. Use the below equation 7 to calculate the Euclidean distance.

\[
\text{Euclidean Distance} = \sqrt{(DT_x - DT_c)^2 + (RT_y - RT_c)^2}
\]  

(7)

In the above equation 7, DT_x and RT_y represent the direct trusts and recommendation trusts nodes in the network respectively and DT_c and RT_c are the randomly selected centroids.

**Step3:** Calculate the distance between the data points (Direct and Recommendation Trust) and Cluster heads (C1, C2, C3).

Therefore,

\[
\begin{align*}
C1 = & (0.3, 0.3) \quad N1 = (0.1, 0.3) \\
C1N1 & \Rightarrow \sqrt{(0.1 - 0.3)^2 + (0.3 - 0.3)^2} = 0.2 \\
C2 = & (0.3, 0.7) \quad N1 = (0.1, 0.3) \\
C2N1 & \Rightarrow \sqrt{(0.1 - 0.3)^2 + (0.3 - 0.7)^2} = 0.4 \\
C3 = & (0.9, 0.4) \quad N1 = (0.1, 0.3) \\
C3N1 & \Rightarrow \sqrt{(0.1 - 0.9)^2 + (0.3 - 0.4)^2} = 0.8
\end{align*}
\]

Similarly, for all other nodes in the network calculate the distance. Table IV denotes the classification of clusters.

**Table IV. Classification of Clusters (Iteration 1)**

<table>
<thead>
<tr>
<th>Nodes</th>
<th>C1 (0.3,0.3)</th>
<th>C2 (0.3,0.7)</th>
<th>C3 (0.9,0.4)</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>0.3</td>
<td>0.2</td>
<td>0.474214</td>
<td>0.806226</td>
</tr>
<tr>
<td>N2</td>
<td>0.2</td>
<td>0.141421</td>
<td>0.309902</td>
<td>0.728011</td>
</tr>
<tr>
<td>N3</td>
<td>0.5</td>
<td>0.538516</td>
<td>0.223607</td>
<td>0.565085</td>
</tr>
<tr>
<td>N4</td>
<td>0.8</td>
<td>0.538516</td>
<td>0.141421</td>
<td>0.14142</td>
</tr>
<tr>
<td>N5</td>
<td>0.3</td>
<td>0.6</td>
<td>0.781025</td>
<td></td>
</tr>
<tr>
<td>N6</td>
<td>0.1</td>
<td>0.808226</td>
<td>0.7</td>
<td>0.316228</td>
</tr>
<tr>
<td>N7</td>
<td>0.3</td>
<td>0.4</td>
<td>0.608276</td>
<td>0.67082</td>
</tr>
<tr>
<td>N8</td>
<td>0.9</td>
<td>0.6</td>
<td>0.680276</td>
<td>0.67082</td>
</tr>
<tr>
<td>N9</td>
<td>0.3</td>
<td>0.7</td>
<td>0.4</td>
<td>0.67082</td>
</tr>
</tbody>
</table>

**Cluster2 \Rightarrow N3(0.5,0.8), N5(0.3,0.9), N9(0.3,0.7)**

**Cluster3 \Rightarrow N4(0.8,0.5), N6(1,0.7), N8(0.9,0.4)**

**Step4:** Calculate the new centroids or header heads by taking the mean of all the calculated data points from each cluster. Therefore,

The new cluster head of cluster 1 after iteration 1 \( \Rightarrow (0.1 + 0.2 + 0.3)/3, (0.3 + 0.2 + 0.3)/3 \Rightarrow (0.2, 0.266)

The new cluster head of cluster 2 after iteration 1 \( \Rightarrow (0.5 + 0.3 + 0.3)/3, (0.8 + 0.9 + 0.7)/3 \Rightarrow (0.66667, 0.8)

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

Now iteration 1 is over. Then repeat the process with the new cluster heads. Table V denotes the new cluster head information and Table VI denotes the classification of clusters after iteration 2.

**Table V. New Cluster Head Information**

<table>
<thead>
<tr>
<th>Cluster</th>
<th>DT</th>
<th>IT</th>
<th>New Cluster Head or Centroids</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>0.2</td>
<td>0.266</td>
<td>(0.2, 0.266)</td>
</tr>
<tr>
<td>C2</td>
<td>0.366667</td>
<td>0.8</td>
<td>(0.366667, 0.8)</td>
</tr>
<tr>
<td>C3</td>
<td>0.9</td>
<td>0.53333</td>
<td>(0.9, 0.53333)</td>
</tr>
</tbody>
</table>

**Table VI. Classification of Clusters (Iteration 2)**

<table>
<thead>
<tr>
<th>Nodes</th>
<th>C1 (0.2,0.266)</th>
<th>C2 (0.3,0.7)</th>
<th>C3 (0.9,0.4)</th>
<th>Cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>0.1</td>
<td>0.050522</td>
<td>0.566667</td>
<td>0.833333</td>
</tr>
<tr>
<td>N2</td>
<td>0.2</td>
<td>0.0666</td>
<td>0.622718</td>
<td>0.775312</td>
</tr>
<tr>
<td>N3</td>
<td>0.5</td>
<td>0.6125</td>
<td>0.133333</td>
<td>0.480742</td>
</tr>
<tr>
<td>N4</td>
<td>0.8</td>
<td>0.644016</td>
<td>0.527046</td>
<td>1.05408</td>
</tr>
<tr>
<td>N5</td>
<td>0.3</td>
<td>0.648183</td>
<td>0.120185</td>
<td>0.703169</td>
</tr>
<tr>
<td>N6</td>
<td>0.1</td>
<td>0.091041</td>
<td>0.641179</td>
<td>0.194368</td>
</tr>
<tr>
<td>N7</td>
<td>0.9</td>
<td>0.105822</td>
<td>0.304425</td>
<td>0.643772</td>
</tr>
<tr>
<td>N8</td>
<td>0.9</td>
<td>0.71271</td>
<td>0.666666</td>
<td>0.133333</td>
</tr>
<tr>
<td>N9</td>
<td>0.3</td>
<td>0.445532</td>
<td>0.120185</td>
<td>0.622719</td>
</tr>
</tbody>
</table>

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

**Cluster1 \Rightarrow N1(0.1,0.3), N2(0.2,0.2), N7(0.3,0.3)**

**Cluster2 \Rightarrow N3(0.5,0.8), N5(0.3,0.9), N9(0.3,0.7)**

**Cluster3 \Rightarrow N4(0.8,0.5), N6(1,0.7), N8(0.9,0.4)**

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

Therefore, The new cluster head of cluster 1 after iteration 2 \( \Rightarrow (0.1 + 0.2 + 0.3)/3, (0.3 + 0.2 + 0.3)/3 \Rightarrow (0.2, 0.266)
The new cluster head of cluster 2 after iteration 2 \( \Rightarrow (0.5 + 0.3 + 0.3)/3, (0.8 + 0.9 + 0.7)/3 \) \( \Rightarrow (0.366667, 0.8) \)

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

The algorithm will stop as two iterations get 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 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 denotes the node’s behavior.

<table>
<thead>
<tr>
<th>Nodes</th>
<th>DT</th>
<th>RT</th>
<th>Iteration 1</th>
<th>Iteration 1</th>
<th>Behaviour</th>
</tr>
</thead>
<tbody>
<tr>
<td>N1</td>
<td>0.1</td>
<td>0.3</td>
<td>C1</td>
<td>C1</td>
<td>Blackhole</td>
</tr>
<tr>
<td>N2</td>
<td>0.2</td>
<td>0.2</td>
<td>C1</td>
<td>C1</td>
<td>Blackhole</td>
</tr>
<tr>
<td>N3</td>
<td>0.5</td>
<td>0.8</td>
<td>C2</td>
<td>C2</td>
<td>Partially Trusted</td>
</tr>
<tr>
<td>N4</td>
<td>0.8</td>
<td>0.5</td>
<td>C3</td>
<td>C3</td>
<td>Trusted</td>
</tr>
<tr>
<td>N5</td>
<td>0.3</td>
<td>0.9</td>
<td>C2</td>
<td>C2</td>
<td>Partially Trusted</td>
</tr>
<tr>
<td>N6</td>
<td>1.7</td>
<td>0.7</td>
<td>C3</td>
<td>C3</td>
<td>Trusted</td>
</tr>
<tr>
<td>N7</td>
<td>0.3</td>
<td>0.3</td>
<td>C1</td>
<td>C1</td>
<td>Blackhole</td>
</tr>
<tr>
<td>N8</td>
<td>0.9</td>
<td>0.4</td>
<td>C3</td>
<td>C3</td>
<td>Trusted</td>
</tr>
<tr>
<td>N9</td>
<td>0.3</td>
<td>0.7</td>
<td>C2</td>
<td>C2</td>
<td>Partially Trusted</td>
</tr>
</tbody>
</table>

After finding the node behavior, now is the time to discuss the tools and techniques used to analyze the results in the next section.

6. RESULTS AND DISCUSSION

The suggested KmeansT has equated with traditional RPL routing protocol and Trust-based RPL in terms of innumerable performance metrics like end-to-end delay, packet delivery and, detection ratio. Assessment of the KmeansT model is steered using the lightweight emulator Cooja on the open-source operating system Contiki 3.0. Additionally, only the Cooja simulator weighs the RPL [22] and Trust-based RPL [27] models. KmeansT uses Tmote Sky (Sensor nodes) motes. In each experiment, we add 10 nodes every 10 rounds. The dataset utilized in the proposed model has been generated using Python. The following Table VIII depicts the simulation parameter settings:

A. Detection Ratio

This parameter plays a crucial role in assessing an algorithm’s efficacy in identifying adversaries, specifically blackhole nodes within a network. In the realm of routing protocols, the conventional RPL protocol lacks inherent detection mechanisms for such malicious nodes, which renders it ineffective in addressing this concern. Consequently, it overlooks the imperative need for adversary detection.

On the contrary, the trust-based variant of the RPL protocol incorporates a singular metric to gauge node trustworthiness. However, its reliance on a solitary metric results in a diminished detection ratio as the prevalence of blackhole nodes escalates.

This weakness stems from the inadequacy of its measurement capabilities when confronted with an amassed number of blackhole nodes.

In contrast, the KmeansT protocol employs a comprehensive approach, integrating both direct trust and recommendation trust metrics to assess node reliability. Moreover, leveraging the K-means clustering algorithm empowers KmeansT to effectively pinpoint blackhole nodes. Consequently, even amidst a rise in the population of black hole nodes, KmeansT exhibits a consistently improving capacity for black hole node detection compared to the trust-based RPL protocol.

Empirical evidence from simulations underscores this disparity in detection capabilities. The average detection ratio of Trust based RPL stands at 19.54%, whereas KmeansT demonstrates a significantly enhanced performance, achieving a detection ratio of 35.38% under conditions where black hole nodes are distributed randomly. Notably, throughout the simulation runs, KmeansT consistently exhibits remarkable efficacy, with its detection ratio peaking at 89.42%. Table IX portrays the uncovering ratio analysis of the proposed KmeansT and Trust-based RPL model.

<table>
<thead>
<tr>
<th>Number of Blackhole Nodes</th>
<th>Trust Based RPL</th>
<th>KmeansT</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>3.5</td>
<td>6.4</td>
</tr>
<tr>
<td>20</td>
<td>12.5</td>
<td>17.8</td>
</tr>
<tr>
<td>30</td>
<td>12.3</td>
<td>27.5</td>
</tr>
<tr>
<td>40</td>
<td>21.3</td>
<td>34.6</td>
</tr>
<tr>
<td>50</td>
<td>27.7</td>
<td>45.6</td>
</tr>
<tr>
<td>60</td>
<td>29.4</td>
<td>54.6</td>
</tr>
<tr>
<td>70</td>
<td>30.1</td>
<td>61.2</td>
</tr>
</tbody>
</table>

Visual representation of these findings is illustrated in Figure 5, depicting a comprehensive analysis of the detection ratios associated with each protocol.

B. Packet Delivery Ratio

The impression of the packet delivery ratio has been scrutinized in Figure 6.
This metric serves as a pivotal indicator of network performance, delineating the proportion of packets successfully reaching their intended destinations relative to those dispatched by the sender node. In pristine network conditions devoid of adversaries, our simulations reveal an impressive packet delivery ratio of approximately 97.5%. However, the introduction of blackhole nodes at regular intervals precipitates a notable decline in this ratio.

Notably, the packet delivery ratio of KmeansT outshines that of the other two protocols, owing to the pivotal role played by the K-means algorithm in blackhole node detection. As the iteration progresses, KmeansT effectively identifies and eliminates these malicious nodes from the network, thereby fostering a higher packet delivery ratio compared to its counterparts. Even in scenarios where 50% of nodes exhibit malicious behavior, KmeansT achieves a commendable 82.5% delivery ratio. Below, Table X illustrates the Packet Delivery Ratio. The corresponding graphical representation is shown in Figure 6.

**TABLE X. Packet Delivery Ratio Analysis**

<table>
<thead>
<tr>
<th>Number of Blackhole Nodes</th>
<th>RPL</th>
<th>Trust Based RPL</th>
<th>KmeansT</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>39.7</td>
<td>52.3</td>
<td>65.3</td>
</tr>
<tr>
<td>20</td>
<td>37.3</td>
<td>48.2</td>
<td>64.3</td>
</tr>
<tr>
<td>30</td>
<td>32.4</td>
<td>45.3</td>
<td>52.3</td>
</tr>
<tr>
<td>40</td>
<td>29.4</td>
<td>34.6</td>
<td>47.4</td>
</tr>
<tr>
<td>50</td>
<td>26.3</td>
<td>32.5</td>
<td>37.4</td>
</tr>
<tr>
<td>60</td>
<td>17.2</td>
<td>19.4</td>
<td>27.4</td>
</tr>
<tr>
<td>70</td>
<td>12.3</td>
<td>17.2</td>
<td>22.5</td>
</tr>
</tbody>
</table>

In contrast, the reliance on a singular metric in Trust-based RPL results in a comparatively lower packet delivery ratio, although it still fares better than traditional RPL. The latter, lacking the capability to detect blackhole nodes, suffers from a diminished packet delivery ratio in the presence of such adversaries.

In essence, the robust detection mechanisms integrated into KmeansT endow it with a superior ability to maintain a high packet delivery ratio, even amidst the manifestation of malicious nodes, thereby underscoring its efficacy in ensuring reliable and efficient network communication.

### C. End-to-end Delay

![Figure 7. End-to-end delay versus Blackhole nodes](https://journal.uob.edu.bh/)

Figure 7 depicts the analysis of end-to-end delay, a critical metric reflecting the time taken for a packet to traverse from its source node to the designated terminus node within the network. Ominously, the KmeansT model reveals a far lower end-to-end delay in evaluation with the other two protocols.

This superior performance of KmeansT can be attributed to its robust mechanisms for blackhole node detection and elimination. By employing multiple trust evaluation mechanisms, KmeansT effectively mitigates the presence of blackhole nodes within the network, thereby minimizing the instances of prolonged delays in packet transmission. Beneath, Table XI depicts the End-to-end analysis and Figure 7 indicates the analogous graph.
In contrast, the trust-based RPL protocol, reliant on a single metric for evaluating node trustworthiness, is more susceptible to the presence of black hole nodes, consequently resulting in higher end-to-end delays compared to KmeansT. However, it still demonstrates better performance in this regard than traditional RPL. Traditional RPL, lacking adequate mechanisms for detecting and addressing blackhole nodes, exhibits the highest end-to-end delays among the three protocols analyzed. The prevalence of black holes within the network exacerbates delays in packet delivery, significantly impacting overall network performance.

Quantitatively, the average end-to-end delay recorded for KmeansT stands at 26.28%, showcasing its efficiency in minimizing delays. In comparison, traditional RPL registers an average delay of 49.07%, while trust-based RPL falls in between with an average delay of 35.2%. These findings underscore the tangible benefits of incorporating robust blackhole detection mechanisms, as exemplified by the KmeansT protocol, in enhancing network efficiency and reducing end-to-end delays.

7. Conclusions

The Internet of Things (IoT) stands as a transformative force, enriching our lives with an array of applications that elevate convenience and comfort. However, this technological paradigm shift brings forth a host of challenges, with security emerging as a paramount concern. Given the delicate nature of information shared within IoT ecosystems, robust security measures are imperative to safeguard against potential threats. Among the myriad security risks, blackhole attacks loom large, poses a grave danger to IoT networks by indiscriminately dropping incoming packets and disrupting essential routing operations. In response to this pressing threat, researchers have introduced the K-Means Clustering-Based Trust (KmeansT) evaluation mechanism as a novel approach to fortify IoT security.

The significance of this study lies in its development of the KmeansT approach, which offers a comprehensive trust evaluation process integrating direct observations and recommendations from network entities. By leveraging the K-Means clustering algorithm, this mechanism enhances trust assessment accuracy, enabling a more nuanced evaluation of node reliability and integrity. Particularly noteworthy is KmeansT’s effectiveness in identifying and mitigating blackhole attacks within IoT environments. Through rigorous mathematical modeling and simulation studies, the study substantiates KmeansT’s efficacy in detecting and neutralizing these threats. Simulation results provide compelling evidence of KmeansT’s superiority in safeguarding IoT networks against blackhole attacks, as evidenced by performance metrics such as end-to-end delay, packet delivery, and detection ratio. In conclusion, the KmeansT approach represents a significant advancement in IoT security, offering a robust framework for safeguarding IoT communication channels against blackhole attacks. By addressing this critical security challenge, KmeansT contributes to the continued evolution and proliferation of IoT technologies, fostering a safer and more secure digital landscape for users worldwide, there remains a necessity to cultivate a more vigorous mechanism to stabilize blackhole nodes.

8. Limitations and Future Directions

Limitations of this research offer further enhancement prospects. First, real-time trust score adaptation to network conditions and performance indicators could improve KmeansT. Adding anomaly detection and encryption to KmeansT could enable multi-layered security against varied attackers. Machine learning for data trend analysis has the potential to improve KmeansT’s threat detection. KmeansT must also be optimized for resource-constrained IoT devices using lightweight versions or methods. Decentralizing trust evaluation among nodes with distributed trust management systems could improve resilience and scalability. Assimilating blockchain technology with KmeansT to record trust scores on a decentralized ledger could boost fidelity and accountability. Finally, KmeansT must be deployed and validated with industry partners in various IoT settings to determine its efficacy and scalability. These advances are crucial to IoT security and KmeansT’s continued development as a powerful threat mitigation solution.

References


https://journal.uob.edu.bh/


Mr. Shameer Mohammed received his MCA from the University of Madras, Chennai. He’s pursuing a PhD in computer science in the Internet of Things at Karpagam Academy of Higher Education in Coimbatore, Tamil Nadu, 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 programs.

Dr. L Gnanaprasanambikai is working as an Assistant Professor in the Department of Computer Science at Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India. She has 17 years of academic experience and 8 years of research experience. She has published various research articles and patents in her academic career. Her current research areas are network security and soft computing.