Privacy-aware secured discrete framework in wireless sensor network

Nandini Sonnappa, Kempanna Muniyegowda
Department of Computer Science and Engineering, Bangalore Institute of Technology, Bangalore, India

Article Info

Article history:

Received Jun 8, 2023
Revised Jul 12, 2023
Accepted Jul 17, 2023

Keywords:

Aggregation
Internet of things
Nodes
Security
Wireless sensor network

ABSTRACT

Rapid expansion of wireless sensor network-internet of things (WSN-IoT) in terms of application and technologies has led to wide research considering efficiency and security aspects. Considering the efficiency approach such as data aggregation along with consensus mechanism has been one of the efficient and secure approaches, however, privacy has been one of major concern and it remains an open issue due to low classification and high misclassification rate. This research work presents the privacy and reliable aware discrete (PRD-aggregation) framework to protect and secure the privacy of the node. It works by initializing the particular variable for each node and defining the threshold; further nodes update their state through the functions, and later consensus is developed among the sensor nodes, which further updates. The novelty of PRD is discretized transmission for efficiency and security. PRD-aggregation offers reliability through efficient termination criteria and avoidance of transmission failure. PRD-aggregation framework is evaluated considering the number of deceptive nodes for securing the node in the network. Furthermore, comparative analysis proves the marginal improvisation in terms of discussed parameter against the existing protocol.

Corresponding Author:

Nandini Sonnappa
Department of Computer Science and Engineering, Bangalore Institute of Technology
Bangalore, India
Email: nandinis@sjcit.ac.in

1. INTRODUCTION

Presently, wireless sensory networks (WSN) are utilized at the forefront of communication systems that include environmental monitoring, smart home automation, and industries. This technology has a widely increasing potential concerning applications of real-time due to its small size, inexpensive and easy deployment [1], [2]. The design of WSNs is dependent on various situations for their application. In applications of the industry as well as automation of the home, in which there is no constraint of energy, delivery packets are the most essential for designing a network. However, in an unsafe environment in which there is a battery that is not replaced or recharged which includes mining, which increased the lifetime of the network, is the main factor in designing the network. Although, the deployed area size has an important role in the design of the WSN. Considering smaller areas, the transmission of the packets is done directly from the sensory nodes to the base station or sinks nodes, whereas, in large areas, the transmission of packets happens via various intermediate nodes to sink nodes [3].

Additionally, the system that is enabled using internet of things (IoT) has an ability for interconnection of various ‘things’ for effective communication as well as sharing of data for one network, there is a wide number of advantages that have attracted various technologists [4], [5]. In the short term, IoT systems are an important part of various sectors that include healthcare, manufacturing, logistics as well as transportation, which enable crucial infrastructure of the IoT [6]. The majority of the intelligent control nodes are linked to the internet in the IoT technology, such nodes are transmitters or sensors that can retrieve data from the other systems and process it without any intervention. Numerous IoT applications are adopted to improve system efficiency and quality in industrial management, transport, and medical; because of the increasing prevalence of the IoT [1]–[3]. IoT technology focuses on multiple activities to achieve the objectives generated through smart services. Figure 1 shows the WSN and IoT integration working. Intelligent devices can interact with the real world through intelligent operations so that consumers may get the proper service whenever and wherever they need it.

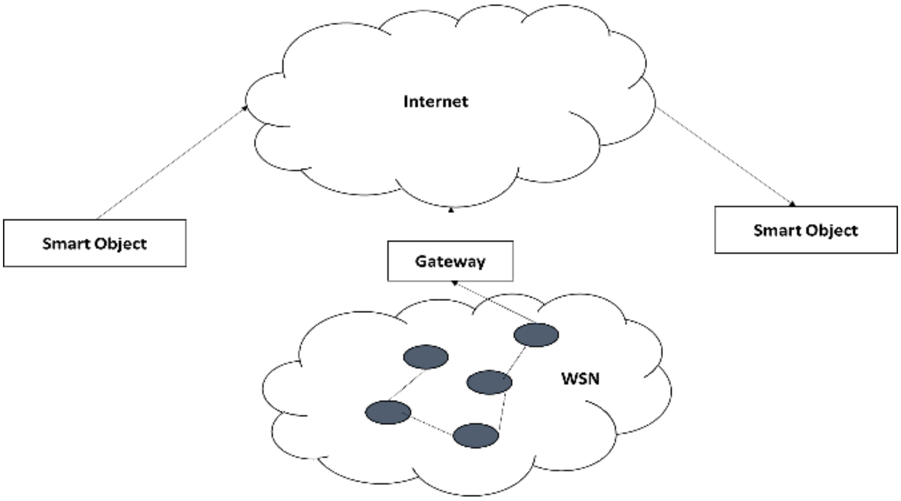

Figure 1. WSN and IoT integration working

Figure 1 shows the integrated architecture of WSN-IoT; it comprises several modules. WSN is where several sensor networks are connected for efficient transmission through the gateway. The other two blocks are smart objects through which data are communicated to humans. Considering the increased complexity in architecture, several uses of heterogeneous devices and threats towards privacy are difficult to be identified, access, and mitigated. However, the large-scale increased complexity of IoT systems introduces a deluge of data. Confidential and sensitive data have been constantly shared among the networks, and privacy, as well as security, are huge concerns that are prevailing in the IoT's crucial infrastructure [7]. Any of the attacks on cyber systems that are vulnerable could compromise the integrity and privacy of huge amounts of data that is sensitive.

The frequency and types of attacks have increased because of substantial technological advancements. Attackers frequently take advantage of the variability in the IoT to create trust challenges and modify behavior to mislead users about the dependability of the sensors and the services offered through them. Moreover, the variety and complexity of IoT systems as well as resource limitations have produced enormous issues in terms of reliability, confidentiality, and transparency, all of which are crucial for the successful deployment of IoT [5]. As a result, the trust evaluation system is used to isolate unreliable objects and identify dishonest behavior. When taking any action, it also overcomes the possibility of ambiguity and minimizes the risks. This can enable IoT infrastructure operation in a controllable environment and prevent unpredictable events and service failures [4]–[6]. Without a reliable security mechanism that prevents the creation of harmful models or at least lessens their influence, IoT systems are unlikely to be widely adopted [6], [7]. As a result, IoT security uses systems for authentication and encryption. Strong authentication and encryption technologies help to address several IoT security concerns. The techniques for secure message transmission between nodes, such as authentication and encryption, serve as the first line of protection against outside threats [8], [9]. These defenses can stop and identify external attacks, but they are unable to deal with insider threats and hostile network nodes. In practice, internal attackers can circumvent these defenses by gaining access to the common key and initiating several attacks against the IoT network. Because of this reason, it is essential to ensure trust to counter internal attacks in the IoT environment. Approaches towards the preservation of privacy have gathered attention over the years due to progress towards information technology that has threatened individual privacy. Although, considering the rise towards the adoption of edge computing as well as fog paradigms for IoT crucial information has led to decreased latency, awareness of location, communication, and data sharing in real-time, and quality of service (QoS) [10], [11]. The devices used for edge computing in IoT crucial information are prone to attacks of privacy [12]. Considering this approach, many techniques for the preservation of privacy that include traditional as well as modern deep learning algorithms are proposed [13]. The existing deep learning methodologies that are implemented based on fog/edge computing for crucial information solutions need expensive computation.

There are various kinds of attacks on IoT crucial information that include attacks for denial of service, Sybil attacks [8], [9]. Access level-based cyber-attacks for IoT crucial networks are classified as passive and active attacks [10]. Active attacks, which are also termed attacks that are security oriented, derange the communication of the network by avoiding the available security protection. On the contrary, passive attacks are termed as attacks that are privacy-oriented such as networks eavesdropping without any disturbance for gaining illicit access to confidential sensitive information. The widely growing IoT crucial information is now prone to various attacks by hackers as well as organized crime syndicates. The growth in the count of threats towards privacy that are targeted towards IoT crucial information results in motivation for the development of various solutions. Although, most of the security approaches that are proposed have the absence of applicability, which may be the result of the complexity of computation, expense, and other factors that are related. Motivated by the above issues, the contribution of the research is given as follows: i) this research takes privacy, security, and reliability into account for aggregation in WSN based IoT environment and develops a privacy and reliable aware discrete (PRD)-aggregation framework; ii) PRD-aggregation framework is an integrated framework that utilizes the designed constraint, discrete, time, and event for preserving the nodes' privacy; furthermore, discrete based mechanism utilizes the threshold setting for nodes for detection at the packet level and node level; and iii) PRD-aggregation is evaluated by inducing the deceptive nodes in the network for detection of dishonest nodes for classification and misclassification of nodes identification.

This research is organized as follows: The first section starts with a background of IoT, sensor nodes, and issues of security and privacy of the nodes and the section concludes with research motivation and contribution. The second section presents the existing approach along with its shortcoming. The third section presents the mathematical modeling of the PR-aggregation model and the fourth section presents the performance evaluation.

2. RELATED WORK

The confidentiality issue of malicious node identification is solved in WSN, in [14] the algorithm is improvised for the neighbor weight trust detection (NWTD) mechanism. This algorithm in parallel updates the trust value to ensure the trust degree for the nodes and fix a minimum peak value in an acceptable range. However, this is capable of ensuring the segmentation of malicious nodes, the main aim of this approach the malicious node detection in WSN's states that in [15] a trusted method is developed based on Dempster-Shafer (D-S) evidence theory is developed by incorporating the indirect and direct trust of third–party nodes to assess the strength of the network. The validation of the data packet, in [16] and [17] based on evaluation have developed a trust model that is focused on the computation of trust degree to handle either direct or indirect levels of trust while taking into account the internal attacks that wireless sensor networks are susceptible. This method minimizes the amount of energy used by the network and generates the trusted benchmark to simplify decision-making, which is easier by frequently updating the degree of trust. To assure network security and dependability, it may further identify between malicious nodes and aging nodes. To reduce the uncertainty of the decisions made by the conventional trust approach, it analyses the signal success rate, node latency, accuracy, and fairness as trust metrics. A multi-attribute trust model was developed in [18] employing fuzzy processing to compute the overall trust value of each node; this demonstrated that the conclusion is accurate. A novel trust management system founded on the D-S evidence theory is discussed herein [19].

The D-S theory was used to analyze the spatiotemporal correlation of data collected by nearby sensor nodes to develop the trust model. Consequently, an evaluation is made to determine which nodes are malicious. In [20], the challenges are addressed by considering the single detection function and failure of the malicious node detection mechanism by introducing a novel malicious node detection model to counter the vicious libel behavior of high-reputation nodes within current WSNs. This model illustrated the Beta Distribution reputation distribution and the implicit reliability of third-party nodes. This guarantees the precise detection of malicious nodes; it additionally incorporates the trust values, which correspond to different attack types. In [20], a methodology of truth discovery for preserving privacy is proposed, although, the client overhead is high. In [21], a two-layer methodology is proposed for fulfilling the requirement of protecting the privacy of the user. In [22], a novel lightweight framework of truth discovery for preserving privacy is proposed that is used for the implementation of two cloud platforms that are non-colluding as well as adopting a homomorphic cryptosystem. In [23], a novel data poisoning disguise attack (DDPA) is proposed against private systems of crowd sensing that are empowered with the methodology of truth discovery. A stealth strategy is proposed that is, the malicious characteristics are disguised for avoiding the detection of methods for truth discovery. Along with this strategy of stealth, the limitation of maximizing the effectiveness of the attack is naturally avoided using optimization problems at the bi-level via structuring that is resolved by an alternate optimization algorithm. In [24], a novel semantic awareness for the preservation of privacy for the trajectory of online location sharing method called Semantic-awareness information-theoretical privacy (SEITP), for the protection of privacy of data as well as semantic during semantic awareness of data is utilized can be preserved completely [25].

3. PROPOSED METHOD

The data that has been gathered from IoT devices must be required for analysis as well as utilization. However, the gathered data could also have information that is sensitive and personal to the user, this hinders the privacy protection of the users. The unwillingness of the users to contribute their information affects the usability of the data. Privacy and reliability while aggregation has been one of the wide research using the consensus-based approach where nodes are required to exchange and disclose their state information to its adjacent node. However, this directly involves the privacy and reliability of nodes. This research work aims to develop a PRD-aggregation mechanism to achieve the trade-off among privacy and aims to guarantee convergence. Moreover, secure data aggregation has been carried out in previous research work; this work utilizes a secured and efficient mechanism and further provides privacy and reliability for the user. Figure 2 presents the proposed PRD model that comprises various modules. At first, the system model is designed based on the connected graph; further total number of variables is selected along with the selection of adjacent nodes. Furthermore, considering the variable, nodes state is updated and constraints are checked and verified.

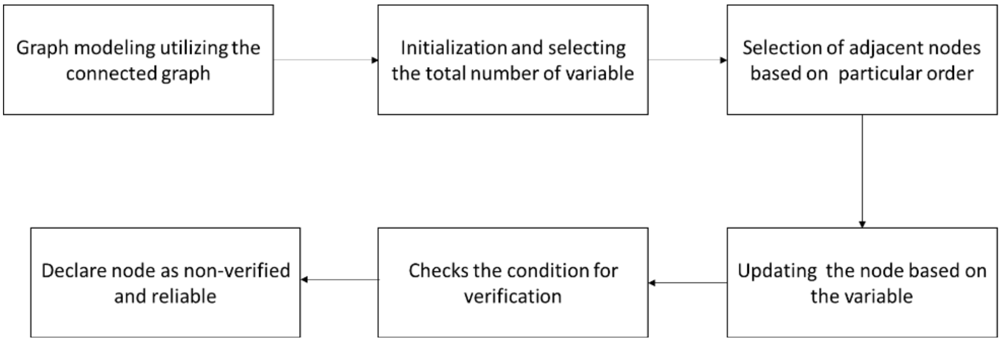
Graph modeling utilizing the connected graph
Initialization and selecting the total number of variable
Selection of adjacent nodes based on particular order
Updating the node based on the variable
Checks the condition for verification
Declare node as non-verified and reliable

Figure 2. Proposed PR-model

3.1. System modeling and problem definition

This section provides the mathematical foundation for modeling among the sensor nodes in the WSN, Let's consider any particular network $Q(Q \geq 2)$ that communicates with adjacent nodes; also, communication is established through the designed directed graph. In a designed graph, $J_g = (Y, H)$, where $Y = [y_1, y_2, y_3, \ldots\ldots y_q]$ and q is the particular node set with G as the edges. Furthermore, we consider particular time step $m \in C \geq 0$ with each node maintaining its particular state. Directed edge between y_m and y_l is denoted through (1).

$$o_{lm} \triangleq (y_l y_m) \in H \tag{1}$$

The equation (1) validates that the sensor node y_m can receive information fromy_l; also, it is assumed that graph designs are connected such that there exists a direct path between the two distinctive nodes. Moreover, the sensor node set which can directly transmit to the node is known as an inner adjacent set of y_m and this can be represented as $Q'_m = [y_l \in Y | (y_l, y_m) \in H]$; also, sensor nodes that can receive direct information from node y_m is known as an outer adjacent set and denoted as Q''_m. Furthermore, the behavior of sensor nodes in the network is discussed. Consider the particular strong connected graph as designed earlier, every node has a certain stage denoted as $\mathrm{b_l}[0]$, thus nodes are required to compute the (2) that satisfies the verifiable and reliable criteria.

$$B' = q(\sum_{o-1}^{q} \; b_o[o])^{-1} \tag{2}$$

The main aim of this research work is to design a particular strategy for a verifiable and reliable strategy for nodes $\mathrm{y_k} \in \mathrm{Y_s}/\mathrm{Y_f}$ while communicating with the nodes. Moreover, the problem is designed to compute the B' as in (1). Moreover, communication takes only among the nodes, which are adjacent to the designed graph, and represents topology. A particular node in set Y could be deceptive and tries to enter the network and violate privacy by trying to violate confidentiality, which means it tries to identify the particular state$\mathrm{b}[0]$. Hence, Privacy can be defined as the node's and network's capability to hide the information, thus proposed model tends to preserve privacy by hiding their confidentiality.

3.2. Discretizing and consensus

Considering the characteristics of data sensing in the WSN-based IoT model, the proposed model is designed such that given time step $\mathrm{m} \in \mathrm{B}_{\geq 0}$, every node holds the characteristics $\mathrm{b_m^v[n], c_m^v[n], c_m^v[n]}$. In here $\mathrm{t^v}$ is considered as yet another variable designed for output. Furthermore, it is assumed that each sensor node is aware of its other adjacent nodes and directly communicates with them; each node gets assigned with a particular order to each of the outgoing edges p_{o_m}. Here $y_n \in Q'_m$, specifically to the link $(\mathrm{y_o}, \mathrm{y_m})$ for node S_{n_m} depicted by S_{o_m}(where $\{S_{o_m} | y_o \in Q'_m\} = \{0,1, \ldots\ldots, G'_m - 1\}$). The proposed model develops the consensus model to process and transmit the information for efficient communication and obtains t^v which is equivalent to the absolute average of the initial states of the node after a certain number of steps. Moreover, it is assumed that every sensor node in the network holds initial states $\mathrm{b_m}[0] \in \mathrm{C}$; furthermore, at every time step $y_m \in Y$ retains the parameters $b_m[n] \in C$ and $c_m[n] \in C$ and state variables $b_m^v[n] \in C$ and $c_m^v[n] \in C$ and $t_n^v[n] = b_m^v[n]/c_m^v[n]$.

$$b_m[n+1] = b_m[n] + \sum_{y_l \in Q''_m} \; 1_{ml}[n] b_l[m] \tag{3}$$

$$c_m[n+1] = c_m[n] + \sum_{y_l \in Q''_m} \; 1_{ml}[n] c_l[n] \tag{4}$$

In (3) and (4), here $1_{ml}[n]$ is zero in case of no communication at the node y_m from the neighbor y_l at iteration m, the following cases are encountered. The proposed model develops the consensus model to process and transmit the information for efficient communication and obtains t^v which is equivalent to the absolute average of the initial states of the node after a certain number of steps. It then transmits $b_m^v[n+1]$, $c_m[n+1]$, to an adjacent $y_n \in Q'_m$ and set the value $\mathrm{b_m[n+1]} = 0$ and $c_b[n+1] = 0$. Two cases are mentioned here.

- First case: $c_m[n+1] > c_m^v[n]$,
- Second case: $c_m[n+1] > c_m^v[n]$ and $b_m[n+1] \geq b_m^v[n]$ is satisfied, the node x_b updates the state variables as (5),

$$c_m^v[n+1] = c_m[n+1], b_m^v[n+1] = b_m[n+1], \; t_m^v[n+1] = \; b_m^v[n+1](c_m[n+1])^{-1} \tag{5}$$

3.3. Designing constraints for privacy, reliability, and verifiability

In this section, we develop a constraint, which aims to preserve the node's information, and any malicious nodes, which aim to obtain information, get removed from the network. Moreover, the main intention of the defined algorithm is to compute the B'. The proposed model follows the event-based approach for security modeling i.e., whenever there is a violation, the event gets triggered. The existing privacy preservation approach utilizes the initial states of the node as $\mathrm{b}'''[0] = \mathrm{b_m}[0] + \mathrm{x_m}$. However, due to the imitation of preservation of privacy, the proposed model introduces the negative variable x_m such that it guarantees the computation of the consensus approach in a certain number of steps. Furthermore, each sensor node holds its privacy value parameter $x_m[n]$, steps O_m, counter variable, and transmission counter variable f_m. The absolute value of the initial value and the number of added variables O_m is required to be larger than the adjacent node y_m. Considering initialization, every node y_m selects the steps and variable parameter to satisfy the given below condition:

Condition 1: Adding steps O_m in the variable of each node y_m required to be larger than or equal to y_m's degree such that the adjacent node receives at least one piece of information.

$$O_m \geq G'_m \tag{6}$$

Condition 2: Moreover, the accumulated variable is infused in computation through the node y_m such that the node state can be computed without any error.

$$x_m = -\sum_{o_m}^{O_m} \ x_m[o_m] \tag{7}$$

Constraint 3: Variable $x_m[o_m]$ is induced to the network by each node based on the event, which needs to be non-negative.

$$x_m[o_m] \geq 0, for\ all\ o_m \epsilon [0, O_m] \tag{8}$$

Constraint 4: Node x_m stops inducing variables such that states can be calculated as (9):

$$x_m[o_m] = 0, for\ all\ o_m \notin [0, O_m] \tag{9}$$

The above constraints indicate that the initial variable is induced in the network hence it is negative and satisfies $u_j \geq G_m''$. Furthermore, while data transmission constraint 1 and constraint 2 hold for each node after certain steps, and in case of violation it will fail to compute the average consensus. The proposed algorithm has a value transfer process in which each node has a connected digraph $H_d = (X, Y)$, which performs executions according to a set of the event-triggered conditions. Each node here $x_b \in X_O$ to ensure privacy in these steps.

- A counter o_m is set to zero and sets the total number of offset-added steps O_m such as $O_m \geq J_m'$ and the set of $(O_m + 1)$ with a positive offset $x_m[o_m] > 0$, where $o_m \in \{0,1,2 \ldots., O_m\}$. The initial negative offset value x_m injects the initial state value $b_l[0]$ to $x_m = -\sum_{om=0}^{O_m} \ y_m[o_m]$.
- To select the $y_o \in Q_m'$ in the order S_{om} to transmit $c_k[0]$ and $b'''[0] = b_m[0] + x_m + x_m[0]$ to the out-neighbor. Then it sets the value to $c'''_m[0] = 0$, $c_m[0] = 0$, and $o_m = o_m + 1$.
- The algorithm is executed, at each step n, node y_m to receive a set of variables $b'''_l[n]$ and $c_l[n]$ for each-in neighbor $y_l \in Q_m'''$. The node y_m updates the variables with $b'''_m[n]$ to check if the events-triggered condition holds. If true then $x_m[o_m]$ to $x_m[n+1]$ and enhances the offset counter o_m by one. It then sets the variables $b_m^v[n+1]$ and $c_m^v[n+1]$ irrespective of $b_m^v[n+1]$ and $c_m^v[\mathrm{P}+1]$. Then it transmits to an out-neighbor $b'''_m[n+1]$ and $c_m[n+1]$ to an out-neighbor in pre-trained order. Here y_m holds the $b_m[n+1]$ and $c_m[n+1]$., No message is received from any of–its neighbors, and with no transmission, the mass variable retains the same.

3.4. PRD-aggregation algorithm

This research work aims to develop a PRD-aggregation mechanism to achieve the trade-off among privacy and aims to guarantee convergence. Moreover, secure data aggregation has been carried out in previous research work. This work utilizes a secured and efficient mechanism and further provides privacy and reliability for the user. Input is taken as the connected graph $J_g = (Y, H)$ with $q = |H|$ edges along with the initial state of $B_l\ [0] \in C$

Step 1: Assigning a particular order R_{o_l} in a given set $\{0,1, \ldots G_m''\}$ for each adjacent node $y_o \in Q_m''$

Step 2: Setting up the counter $f_m = 0$ along with index (priority based) h_m to f_m

Step 3: Setting up the counter $o_m = 0$, selects $O_m \in C > 0$ where $O_m \geq \ G_m''$

Step 4: Setting up $B_m' = \ B_m[0] + x_m, C_m[0] = 1, C_m^v[0] = 1\ \&\ C_m^v[0] = B_m'[0]$

Step 5: Choosing adjacent node $y_o \in \ Q_m''$ such that $R_{om} = h_m$ and transmit $C_m[0]$ and $C_m' + x_m[0]$ to a particular adjacent node. Furthermore, setting $B_m'[0] = 0$, $C_m'[0] = 0$, $n_m = n_m + 1$

Step 6: Setting $f_m = f_m + 1$ and $h_m = f_m\ mod\ O_m''$

Step 7: Considering the iteration of $n = 0,1,2 \ldots.\ every$ node $\mathrm{y_m}$ carries out the following operation.

Step 8: If it receives $\mathrm{B_l'[N], C_l[N]}$ from an adjacent node $\mathrm{y_m} \ \in \ \mathrm{Q_m'''}$ and updation is carried out with (1) and (2).

Step 9: If (1) and (2) hold then transmit the information about nodes for preserving privacy.

Step 10: Output as $\mathrm{t_m^v[n]}$ for each node $y_m \in Y$.

In the above algorithm, the first o_m to zero and select the total number of variables adding with O_m steps such that $O_m \geq G_m''$ and with $(O_m + 1)$ positive variable $x_m[o_m] > 0$. Furthermore, considering $y_o \in Q_m'$ following order S_{om} and transmit $B_m'[0]$ and $B_m'[0]$. Furthermore, while executing the algorithm, at every step m, node y_m receive a set of requests for packet transmission from each adjacent node, conditions are checked for verification, and if holds then the data packets are transmitted and the node remains else nodes are discarded out of the network.

4. PERFORMANCE EVALUATION

While performing the aggregation it is very important, that reliable and truthful data be collected. PRD-aggregation aims to assure the privacy of sensor nodes as well the reliability of the data. Moreover, PRD-aggregation is evaluated considering the dishonest sensor nodes. It is evaluated on the system configuration that includes 2 TB of the hard disk loaded with 16 GB of random-access memory (RAM) along with 2 GB NVidia Cuda-enabled graphics. The proposed model here works on evaluating an incorrect identification of the node that results in unevenness in the network by accommodating various parameters like identification of the correct node, identification of the wrong node, and computing the throughput for 30, 40, and 50 nodes. Additionally, a comparative analysis is done between the proposed model with the existing model to ensure the model's security and efficiency and conclude that our proposed system performs better than the existing system.

4.1. Energy utilization

Figure 3 shows the utilization of energy over the deceptive nodes, the graph is plotted for energy utilization by the deceptive nodes for 30, 40, and 50 nodes. For 30 nodes there is a steep increase in the after 4.3 value, whereas for 40 nodes there is a steep rise after 4.5 value. For 50 nodes, there exists a steep rise after 4.9 value. An overall comparison is shown below for 30 40 and 50 nodes.

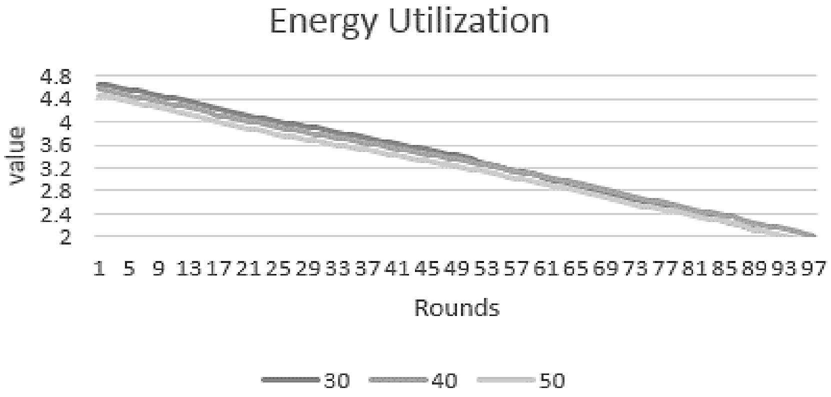

Figure 3. Energy utilization of deceptive nodes

4.2. Correct identification

Identification of node and packet marks the evaluation and performance of the PRD-Aggregation model. This section evaluates the security framework based on packet level and node level. The correct nodes are identified wherein a comparison is made between the existing system and the proposed system by evaluating the correct identification of nodes with 30, 40, and 50 nodes.

4.3. Packet level

In this section, the packets are identified wherein a comparison is made between the existing system and the proposed system by evaluating the comparison with 30, 40, and 50 nodes. Figure 4 shows the comparison of the stated above; in the context of 30, 40, and 50 compromised nodes, the existing system detects 70, 76, and 77 sensor nodes respectively. In the context of 30, 40, and 50 compromised nodes, the proposed model identifies 89, 97, and 96 nodes respectively.

4.4. Node level

In this section, the correct nodes are identified wherein a comparison is made between the existing system and the proposed system by evaluating the correct identification of nodes with 30, 40, and 50 nodes. Figure 5 shows the comparison of the stated above; in the context of 30, 40, and 50 compromised nodes, the existing system detects 41.75 and 66 sensor nodes respectively. In the context of 30, 40, and 50 compromised nodes, the proposed model identifies 97.99 and 97 nodes respectively.

4.5. Wrong node identification

Figure 6 depicts the wrong identification of nodes for 30, 40, and 50 sensor nodes. In 30 nodes context, the existing model identifies 34 wrong nodes whereas the proposed model wrongly identifies 4 nodes. In 40 nodes, the existing model wrongly identifies 36 nodes whereas the proposed model wrongly identifies 2 nodes. In 50 nodes context, the existing model wrongly identifies 35 wrong nodes whereas the proposed model wrongly identifies 4 nodes.

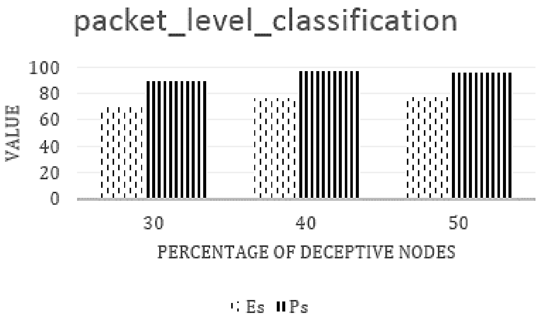

Figure 4. Packet level classification

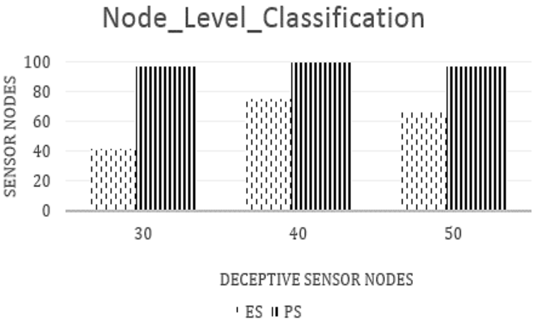

Figure 5. Correct node identification

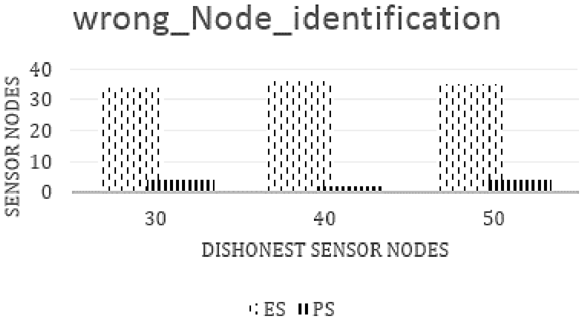

Figure 6. Wrong node identification

4.6. Throughput

Throughput is defined as the amount of work done in a specific amount of time; it displays the models' efficiency. This section evaluates the throughput parameter at the packet level and node level. *Throughput_packet_level* is identified wherein a comparison is made between the existing system and the proposed system.

4.7. Packet level

In this section, the *throughput_packet_level* is identified wherein a comparison is made between the existing system and the proposed system by evaluating the comparison with 30, 40, and 50 nodes. Figure 7 shows the comparison of the stated above; in the context of 30 compromised nodes, the existing system detects 0.266 *packet_level* and the proposed model identifies 0.3382 *packet_level*. Consequently, in the context of 40 sensor nodes, the existing system identifies 0.2128 *packet_level* whereas the proposed model identifies 0.2716 *packet_level*. For 50 sensor nodes, the existing system identifies 0.1155 *packet_level* whereas the proposed model identifies 0.144 *packet_level*.

4.8. Node level

Figure 8 shows the throughput comparison for 30, 40, and 50 sensor nodes. In the case of 30 compromised nodes, the throughput of the existing model is 0.123. For the proposed model, it is 1.09933. In the case of 40 compromised nodes, the throughput of the existing model is 0.3 and for the proposed model, it is 0.891. In the case of 50 compromised nodes, the throughput of the existing model is 0.33 and for the proposed model, it is 0.693.

4.9. Comparative analysis

PRD-aggregation framework is evaluated in the previous section and analyzed in this section. Comparative analysis is carried out by analyzing the performance enhancement over the low-energy adaptive clustering hierarchy (LEACH) protocol. Several parameters are mentioned for the comparative analysis.

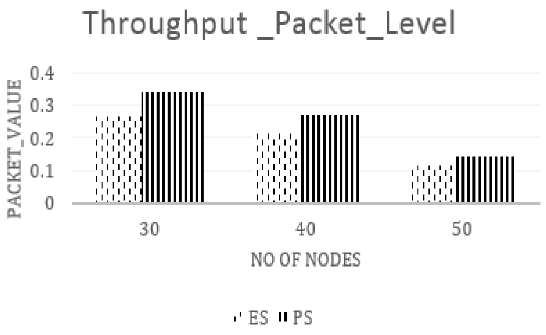

Figure 7. Throughput *packet_level* comparison

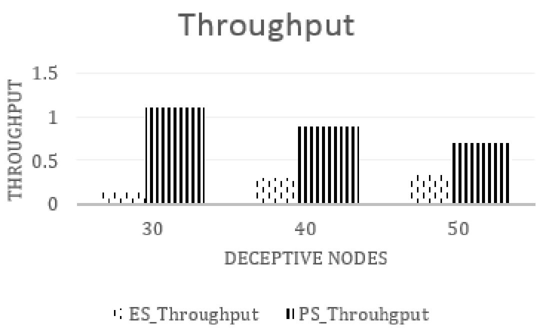

Figure 8. Throughput comparison

4.10. Improvisation over correctly classified nodes

This section displays the comparative analysis and shows the percentage improvisation for the proposed model from the existing model. Table 1 shows the deceptive nodes improvisation table. The improvisation is carried out on correct node identification for 30 nodes is 8,159%. The improvisation for 50 nodes the improvisation is 2.758% for 50 nodes the improvisation is 3,803%.

4.11. Improvisation over wrong classified nodes

Table 2 shows the wrong node identification. For wrong node identification, the improvisation for 30 nodes is 1.25%. The improvisation for 40 nodes is 1.05263%. The improvisation for 50 nodes is 15.89%. Above mentioned table shows the improvisation.

Table 1. Deceptive nodes improvisation table

Deceptive nodes improvisation table	
Deceptive nodes	Improvisation
30	8.159
40	2.758
50	15.78

Table 2. Shows the wrong node identification

Wrong node identification	
Wrong node identification	Improvisation
30	1.25%
40	1.05263%
50	15.8%

4.12. Throughput improvisation

For *throughput_packet_level*, the improvisation for 30 nodes is 2.38%, and for 40 nodes, the improvisation is 2.427%. For 50 nodes, it is 2.196%. For throughput at the node level. Table 3 shows the throughput performance. The improvisation for 30 nodes is 15.95%, for 40 nodes, the improvisation is 9.924%, and for 50 nodes, it is 7.096%.

Table 3. Throughput performance

Deceptive nodes	Packet level (improvisation in percentage)	Node level (Improvisation in percentage)
30	2.38	15.95
40	2.42	9.92
50	2.19	7.096

5. CONCLUSION

Privacy has been one of the major concerns in data transmission, and aggregation due to the adoption of a consensus-based mechanism for efficient operation; this research work develops a novel mechanism PRD-aggregation mechanism for ensuring the sensor nodes' privacy to enhance the confidentiality. The PRD-aggregation novelty lies in the optimization of the consensus approach through the proposed algorithm. PRD-aggregation framework is evaluated at node and packet level for security considering 30, 40, and 50 deceptive nodes. Moreover, the evaluation shows the significant improvisation over the existing leach protocol with classification, misclassification, and throughput. In terms of throughput, the PRD-aggregation framework achieves improvisation of 2-2.5 at the packet level whereas near or more than 10% at the node level. The future scope of research lies in the adoption of data integrity technologies such as blockchain due to the rise of deep learning-based attack models.

ACKNOWLEDGMENT

I would like to express our sincere gratitude to all those who have supported and contributed to this research project. Primarily, I extend our heartfelt thanks to our guide for his unwavering guidance, invaluable insights, and encouragement throughout the research process. No funding is raised for this research.

REFERENCES

[1] I. Lee and K. Lee, "The internet of things (IoT): applications, investments, and challenges for enterprises," *Business Horizons*, vol. 58, no. 4, pp. 431–440, Jul. 2015, doi: 10.1016/j.bushor.2015.03.008.

[2] N. Ma, H. Zhang, H. Hu, and Y. Qin, "ESCVAD: an energy-saving routing protocol based on voronoi adaptive clustering for wireless sensor networks," *IEEE Internet of Things Journal*, vol. 9, no. 11, pp. 9071–9085, Jun. 2022, doi: 10.1109/JIOT.2021.3120744.

[3] W. Zhou, Y. Jia, A. Peng, Y. Zhang, and P. Liu, "The effect of IoT new features on security and privacy: New threats, existing solutions, and challenges yet to be solved," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1606–1616, Apr. 2019, doi: 10.1109/JIOT.2018.2847733.

[4] H. Attaullah *et al.*, "Fuzzy-logic-based privacy-aware dynamic release of IoT-enabled healthcare data," *IEEE Internet of Things Journal*, vol. 9, no. 6, pp. 4411–4420, Mar. 2022, doi: 10.1109/JIOT.2021.3103939.

[5] M. A. Husnoo, A. Anwar, R. K. Chakrabortty, R. Doss, and M. J. Ryan, "Differential privacy for IoT-enabled critical infrastructure: a comprehensive survey," *IEEE Access*, vol. 9, pp. 153276–153304, 2021, doi: 10.1109/ACCESS.2021.3124309.

[6] S. T. Mehedi, A. Anwar, Z. Rahman, K. Ahmed, and R. Islam, "Dependable intrusion detection system for IoT: a deep transfer learning based approach," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 1006–1017, Jan. 2023, doi: 10.1109/TII.2022.3164770.

[7] Y. Sei and A. Ohsuga, "Private true data mining: differential privacy featuring errors to manage internet-of-things data," *IEEE Access*, vol. 10, pp. 8738–8757, 2022, doi: 10.1109/ACCESS.2022.3143813.

[8] J. A. Onesimu, J. Karthikeyan, J. Eunice, M. Pomplun, and H. Dang, "Privacy preserving attribute-focused anonymization scheme for healthcare data publishing," *IEEE Access*, vol. 10, pp. 86979–86997, 2022, doi: 10.1109/ACCESS.2022.3199433.

[9] J. Andrew, J. Karthikeyan, and J. Jebastin, "Privacy preserving big data publication on cloud using mondrian anonymization techniques and deep neural networks," in *2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS)*, Mar. 2019, pp. 722–727, doi: 10.1109/ICACCS.2019.8728384.

[10] A. V. Dastjerdi and R. Buyya, "Fog computing: helping the internet of things realize its potential," *Computer*, vol. 49, no. 8, pp. 112–116, Aug. 2016, doi: 10.1109/MC.2016.245.

[11] R. Deng, R. Lu, C. Lai, and T. H. Luan, "Towards power consumption-delay tradeoff by workload allocation in cloud-fog computing," in *2015 IEEE International Conference on Communications (ICC)*, Jun. 2015, pp. 3909–3914, doi: 10.1109/ICC.2015.7248934.

[12] M. Mukherjee *et al.*, "Security and privacy in fog computing: challenges," *IEEE Access*, vol. 5, pp. 19293–19304, 2017, doi: 10.1109/ACCESS.2017.2749422.

[13] R. Lu, K. Heung, A. H. Lashkari, and A. A. Ghorbani, "A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT," *IEEE Access*, vol. 5, pp. 3302–3312, 2017, doi: 10.1109/ACCESS.2017.2677520.

[14] F. Zawaideh, M. Salamah, and H. Al-Bahadili, "A fair trust-based malicious node detection and isolation scheme for WSNs," in *2017 2nd International Conference on the Applications of Information Technology in Developing Renewable Energy Processes & Systems (IT-DREPS)*, Dec. 2017, pp. 1–6, doi: 10.1109/IT-DREPS.2017.8277813.

[15] N. Hrovatin, A. Tošić, M. Mrissa, and J. Vičič, "A general purpose data and query privacy preserving protocol for wireless sensor networks," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 4883–4898, 2023, doi: 10.1109/TIFS.2023.3300524.

[16] L. Shi, W. X. Zheng, Q. Liu, Y. Liu, and J. Shao, "Privacy-preserving distributed iterative localization for wireless sensor networks," *IEEE Transactions on Industrial Electronics*, vol. 70, no. 11, pp. 11628–11638, Nov. 2023, doi: 10.1109/TIE.2022.3231272.

[17] V. Ram Prabha and P. Latha, "Fuzzy trust protocol for malicious node detection in wireless sensor networks," *Wireless Personal Communications*, vol. 94, no. 4, pp. 2549–2559, Jun. 2017, doi: 10.1007/s11277-016-3666-1.

[18] W. Zhang, S. Zhu, J. Tang, and N. Xiong, "A novel trust management scheme based on Dempster–Shafer evidence theory for malicious nodes detection in wireless sensor networks," *The Journal of Supercomputing*, vol. 74, no. 4, pp. 1779–1801, Apr. 2018, doi: 10.1007/s11227-017-2150-3.

[19] G. Yang, G. S. Yin, W. Yang, and D. M. Zuo, "A reputation-based model for malicious node detection in WSNs," *Harbin Gongye Daxue Xuebao/Journal of Harbin Institute of Technology*, vol. 41, no. 10, pp. 158–162, 2009.

[20] C. Miao *et al.*, "Privacy-preserving truth discovery in crowd sensing systems," *ACM Transactions on Sensor Networks*, vol. 15, no. 1, pp. 1–32, Feb. 2019, doi: 10.1145/3277505.

[21] Y. Li *et al.*, "An efficient two-layer mechanism for privacy-preserving truth discovery," in *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Jul. 2018, pp. 1705–1714, doi: 10.1145/3219819.3219998.

[22] C. Miao, L. Su, W. Jiang, Y. Li, and M. Tian, "A lightweight privacy-preserving truth discovery framework for mobile crowd sensing systems," in *IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*, May 2017, pp. 1–9, doi: 10.1109/INFOCOM.2017.8057114.

[23] Z. Li, Z. Zheng, S. Guo, B. Guo, F. Xiao, and K. Ren, "Disguised as privacy: data poisoning attacks against differentially private crowdsensing systems," *IEEE Transactions on Mobile Computing*, vol. 22, no. 9, 2022, doi: 10.1109/TMC.2022.3173642.

[24] Z. Zheng, Z. Li, H. Jiang, L. Y. Zhang, and D. Tu, "Semantic-aware privacy-preserving online location trajectory data sharing," *IEEE Transactions on Information Forensics and Security*, vol. 17, pp. 2256–2271, 2022, doi: 10.1109/TIFS.2022.3181855.

[25] J. N. Al-Karaki and G. A. Al-Mashaqbeh, "SENSORIA: a new simulation platform for wireless sensor networks," in *2007 International Conference on Sensor Technologies and Applications (SENSORCOMM 2007)*, Oct. 2007, pp. 424–429, doi: 10.1109/SENSORCOMM.2007.4394958.

BIOGRAPHIES OF AUTHORS

Nandini Sonnappa received the B.Eng. degree in information science and engineering from Visvesvaraya Technological University, Belagaum, and Karnataka state in 2008 and the M.Tech. Degree in computer science and engineering from Visvesvaraya Technological University, Belagaum, Karnataka state in 2011 and pursuing Ph.D. in Visvesvaraya Technological University, Belagaum, and Karnataka. Currently, she is an Assistant Professor at the Department of Information Science and Engineering, S J C Institute of Technology, Visveswaraya technological University. Her research interests include data communication, data structures and applications, cryptography and network security, cyber security, cloud computing, and IoT. She can be contacted at email: nandinis@sjcit.ac.in.

Kempanna Muniyegowda completed Ph.D. in Karpagam University in the year 2016 in the field of network security recognized as research guide under VTU University NA students are awarded Ph.D. and Three number of students are pursuing Ph.D. under my guidance. Major areas of research interest are network security and web security. Published 09 number of papers in recognized journals/conferences. His research interests include data communication, cryptography and network security, cyber security, cloud computing, and IoT. lifetime Member in International Association of Engineers (IAENG), Member Number: 269338. He can be contacted at email: kempsindia@gmail.com and kempannam@bit-bangalore.edu.in.