EPO Based Clustering and Secure Trust-Based Enhanced LEACH Routing in WSN

L. Rajesh and H. S. Mohan

Abstract *Background*: In wireless sensor network (WSN), one of the best method is data aggregation for energy conservation. In numerous WSN applications, privacy preserving has become a very important issues with the development of WSNs. *Purpose*: Sensors are vulnerable for security threats due to the open deployment. Moreover, most of the routing algorithms does not consider security. Various methods like trust management, key management, firewalls, and intrusion detection for providing security were considered in the existing work. Between them, enhanced security is given by the trust management. Thus, in this paper, we used the secure trust-based enhanced LEACH routing (STELR) protocol to resolve security concerns. *Methods*: Here, cluster head is selected by using emperor penguin optimization (EPO) algorithm for data aggregation, and STELR is used to enable the security for routing. We can enable secure communication by checking the trustworthiness of a node. *Results*: The work is implemented in MATLAB platform, and the results are measured in terms of packet delivery ratio (PDR), throughput, end-to-end (E2E) delay, detection accuracy, energy dissipated, and trust probability (TP). *Conclusions*: Finally, the security concerns are resolved, which shows proposed approach that gives better performance when compared to existing techniques like LEACH and SDILR protocols.

Keywords Wireless sensor network · Emperor penguin optimization · LEACH protocol · Trust value · Cluster head · Key signature generation · Key signature verification data aggregation and routing

1 Introduction

Numerous applications are in wireless sensor network for monitoring various environments. It includes less cost and small sensing devices along with wireless radio transceiver [1]. WSN contain hundreds or thousands of sensor nodes [2]. The sensor

L. Rajesh (✉) · H. S. Mohan
Department of Information Science and Engineering, SJB Institute of Technology, Bengaluru 560060, India

P. Karrupusamy et al. (eds.), *Sustainable Communication Networks and Application*, Lecture Notes on Data Engineering and Communications Technologies 93,
https://doi.org/10.1007/978-981-16-6605-6_3

senses and collects the data which is transmitted to one or more than one sink nodes. The ability of sensor is to communicate directly or each other to an external base station [3]. This network does not need any external supply for data gathering which is one of the main advantages [4]. The WSN main applications are military surveillance, building safety monitoring, forest fire detection, wild habitat monitoring, etc. [5].

In remote sensing application, limited battery power is the key challenges [6]. To reduce energy consumption, one of the main solution is cluster-based data aggregation which deals numerous copies of data and incorporate more efficient information by reducing data packets [7, 8]. One of the fundamental process is data aggregation which can be consider to diminish communication overhead and energy consumption in WSNs for save the limited resources [9].

Nowadays, Internet of Things (IoT) is one of the part of WSNs. The self and openness organization characteristics expose its weakness to the attackers that lead to loss worse damage and loss original construction [10, 11]. Feasible and efficient secure data aggregation construct a basis for WSNs applications in various filed like economic, political, military, etc. [12]. To protect WSNs from hostile attacks, a lot of efforts have been made by the researchers. However, the existing methods have lower privacy preserving capability, high communication cost, and large computational, and also few technical challenges need to solve like secure data aggregation, efficiency, and energy in WSNs [13].

The key contribution is given below,

- Placement of sensor node in the network for the secure communication without eavesdropping attack.
- Emperor penguin optimization algorithm is used for selecting the cluster head based on energy and distance.
- Secure trust-based enhanced LEACH routing (STELR) is used to enable the security for routing.

Organization: The related work is given in Sects. 2 and 3 which gives network model, thread model, emperor penguin optimized cluster head selection, and STELR. Sect. 4 involves the performance of our work. The conclusion is given in Sect. 5.

2 Related Work

Data security and data aggregation issues was addressed by Ranjani et al. [14] for WSN of IoT application. To give secure data transmission, an energy-efficient cluster-based data aggregation (ECBDA) method was used. Bayesian fusion algorithm and data aggregation was performed by cluster head to enable the security. Secure communication was enabled by checking the node trustworthiness. Based on the behavior of node, the algorithm of Bayesian fusion calculates the trust. The parameters like network residual energy, energy consumption, delay, and throughput was mainly considered.

Table 1 Analysis of existing comparison

Author	Method	Jitter	Network lifetime (ms)	Delay
Ranjani et al. [14]	Efficient sink selection scheme (ES3)	0.05 s	90	1.5 s
Kumar et al. [15]	SDARP protocol	4 s	88	8 s
Sahu et al. [16]	Token-based data security protocol	–	85	–
Fu et al. [4]	DCA-SF protocol	–	75	–

Data security and network traffic issues were concentrated by Kumar et al. [15], and in order to diminish these issues high data gathering, a security-based data-aware routing protocol was implemented for adhoc sensor network application. Two phases were conducted in the security model. Optimal cluster head was used in the first phase to monitor the cluster member and CH behavior. This model was enhanced in the second phase with the data gathering algorithm. The parameters like energy efficiency, network lifetime, and end-to-end delay was calculated.

In case of inter-cluster communication, the data security operation was performed by Sahu et al. [16] which depends on token cluster identification for WSN application. Before communication was initiated, the cluster sender checks the identifications of the receiver cluster. By cluster head node, each cluster was represented. The head node identification number was called as the token number. Parameters like initial energy, sensing range, and number of cluster and cluster head were calculated.

For detecting a selective forwarding attack, a data clustering algorithm was proposed by Fu et al. [4] for WSN application. By clustering their cumulative forwarding speeds, the malicious cluster heads were captured and isolated. The algorithm has been improved by altering the parameters of the data clustering algorithm. The parameters like high detection speed ratio, low false detection rate, and missed detection rate was calculated.

To ameliorate the traffic flow, Bestak [17] modeled an intelligent traffic control device using adhoc network. In the cloud, the smart resource management for massive data processing system was developed by Chandy [18]. The parameters like network utilization, disk utilization, memory utilization, and CPU utilization was measured and considered. The performance like jitter, network lifetime, and delay is compared with several existing techniques, which is given in Table 1.

3 Proposed Methodology

In this paper, we addressed the security issues by using STELR protocol. Initially, the sensor node is placed in the network for secure communication without attacking eavesdropping. Here, cluster head is selected by using EPO algorithm based on distance and energy for data aggregation, and STELR is used to enable the security

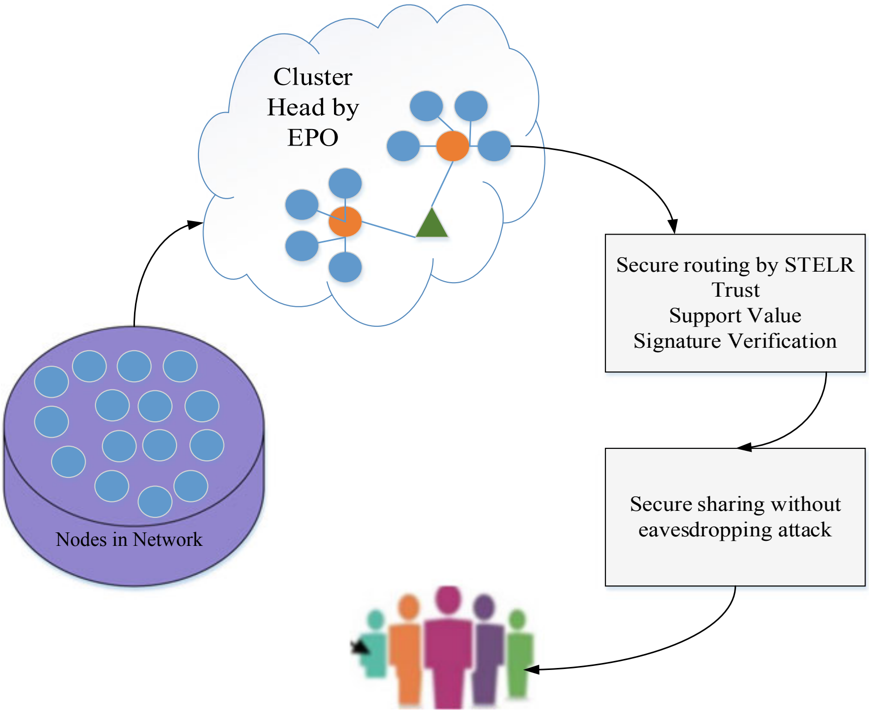


Fig. 1 Procedure for proposed work

for routing. We can enable secure communication by checking the trustworthiness of a node. To ensure the avoidance of eavesdropping, support value-based trust authentication is given to the primary users after the cluster formation in the proposed STELR protocol. The support value-based trust authentication will verify the trust value, key signature generation value, and key signature verification value. The above flow of work and working procedure is given in Fig. 1.

3.1 Network Model

In WSN, the sensor network is modeled as $G\,(V, E)$ which is connected graph, where set of sensor node is represented as V, number of sensor node is denoted as $|V| = N$, and set of wireless link connecting to the sensor node is represented as E. Within transmission range, the sensor nodes needs to be communicated, and for that, every sensor node is fitted out along with wireless transceiver.

Network data aggregation process, the sensor nodes can be allocated in to three groups like sink, aggregation, and ordinary node. The aggregation node and ordinary node is same; however, both tasks are not similar. The final aggregation results is received by sink node and forwarding the query request.

The function of data aggregation is represented as $Y_1(t) = f_1(d_1(t) + d_2(t) + ... + d_n(t))$, where the individual sensor reading for node i at time instant t is represented as $d_i(t) = (i = 1, 2, 3..., N)$. Typically, the aggregation function is a sum of count, min, max, average, sum, etc. Nowadays, most of the researchers have only focused on the sum function. Some statistical function like standard deviation, average, and count are all based on sum (). To estimate by the sum function, the nonlinear function like min () and max () can also be used, and then, it will be forwarded to the sink, and in result, more energy will be saved.

3.2 Threat Model

In the field of the WSNs, the security issues have become more and more important. Different attacks can be launched by the hacker to undermine the data security. In WSN, to protect data privacy, we are mostly attentive on avoiding eavesdropping attacks.

Threats from untrusted eavesdropping listening or intercepting to packets is included by the threat model. The attacker had the capability to learn all communications in attack processing and by monitoring the wireless link to obtain user privacy information. Let us assume, the security mechanism is adopted in WSN to access the attackers by capturing a normal node. Data privacy will be threatened, if the attacker captures the privacy node data than the other nodes.

3.3 Emperor Penguin Optimized Cluster Head Selection

The cluster head for the cluster member is selected in the cluster head selection phase, having a minimum distance and the greatest residual energy. Here, by using fitness of EPO, the cluster head is selected.

3.3.1 Cluster Boundary Determination

The sensor node is chosen to establish the cluster's boundary. To define the cluster boundary, the Euclidean distance is used. The boundary requirement is less than the Euclidean distance.

$$\Psi = \Delta\Phi \tag{1}$$

where Ψ denotes the cluster boundary and Φ denotes the Euclidean distance

$$\text{Fitness Function } = \text{CH} = \sum_{\max} E + \sum_{\min} D \tag{2}$$

where E represents the energy of nodes and D represents the distance between the nodes.

3.3.2 Energy Calculation Around Boundary

The sensor node's energy between the boundary positions is determined. To model quantitatively, we assume the energy while the power is on and when the power is off. The energy is in charge of sensor nodes in various areas. The energy calculation around the boundary is as follows:

$$E^{'} = \left(T - \frac{\text{Max}_{\text{iter}}}{x - \text{Max}_{\text{iter}}}\right) \tag{3}$$

$$E = \begin{cases} 0, \text{ if } P > 1 \\ 1, \text{ if } P < 1 \end{cases} \tag{4}$$

where current iteration is defined as x, maximum number of iteration is denoted as Max_{iter}, time is denoted as T, and power is denoted as P.

3.3.3 Distance Between Sensor Nodes

Following the construction of the boundary, the distance between the sensor nodes and the best optimal solution are computed. The other search agent will update their location based on the current best optimal solution. The following is the distance between the sensor nodes:

$$\vec{D}_{\text{ep}} = \text{Abs}\left(S\left(\vec{A}\right).\vec{P}(y) - \vec{E}.\vec{P}_{\text{ep}}(y)\right) \tag{5}$$

In which the distance between SNs are denoted as $\vec{D}_{\text{ep}}$, the force of node which move toward the best optimal solution is represented S, best optimal solution is defined as $\vec{P}$, current iteration is indicated as y, $\vec{A}$ and $\vec{E}$ are utilized to avoid the collision, the position of sensor node is denoted as $\vec{P}_{\text{ep}}$.

3.3.4 Optimal Solution

The cluster head is chosen based on sensor node distance and energy. The cluster heads are updated based on the best optimal solution. This is in charge of choosing the cluster head for each cluster in a given space. The following equation is used to update the CH for each cluster in an SN.

$$\vec{P}_{\mathrm{ep}}(y+1) = \vec{P}(y) - \vec{A} \cdot \vec{D}_{\mathrm{ep}} \tag{6}$$

The updated CH is represented as $\vec{P}_{\mathrm{ep}}(y+1)$, and updated equation is compared with the fitness Eq. (5), if it satisfied the fitness function, the CH is selected with optimal solution or else the condition will be continued. The proposed cluster formation reduces the energy consumption because only the cluster head nodes are used for the transmission. Here, the secondary user accesses the data over time and is referred to as an eavesdropping attack. In order to avoid this, support value-based authentication is given to the data transmission node.

3.4 *Secure Trust Based Enhanced LEACH Routing (STELR)*

The support value (SV)-based trust authentication (TA) (SV-TA) is provided to main users in the proposed STELR protocol to avoid eavesdropping attacks. When the information is transmitted, the security is given by the trust authentication on the primary user nodes by creating support value based signature verification. This will not help to get the malicious nodes (MN) and got the access of primary user nodes. The following steps are trust, key signature creation, and verification stages.

3.4.1 Trust Based Data and Communication

Based on communication and data trust, the trust value will be calculated. By the beta distribution factor, the communication trust is obtained. Beta probability density function can be used to determine the probability of all binary events. By using beta density function, the communication trust is defined.

$$f(y_i; \alpha_i, \beta_i) = \Gamma(\alpha_i + \beta_i) \big/ \Gamma\alpha_i \Gamma\beta_i y_i^{\alpha_i - 1}(1 - y_i)^{\beta_i - 1} \tag{7}$$

where number of successful transmission is denoted as α_i, number of cluster member failure transmission cm_i is denoted as β_i.

Data trust is calculated by using the Gaussian distribution function.

$$f(y_i; M, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(y_i - M)^2/2\sigma^2} \tag{8}$$

$$M = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{9}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - M)^2} \tag{10}$$

where variance and mean of the sensed data is represented as σ and M from the sensors group in a cluster. The sensed value of cm_i is represented as y_i.

3.4.2 Key Signature Generation

Before sending the information, the key signature is generated. The key is necessary for the primary user to access the information. If the data is not accessed by the user, then the node is attacked. After new node is searched, it depends upon the key access to transfer the data. In subsequent stages, the key signature generation process is described.

Stage 1: The random number of group public key $g_{\text{pk}} = (g_1, g_2)$ and message $N \in \{0, 1\}^*$ is the input for the key signature, i.e., $\text{Sign}\left(g_{\text{pk}}, N\right)$.
Stage 2: The random numbers δ and ϕ are selected.
Stage 3: By using Eqs. (11) and (12), the helper values B_1 and B_2 are computed,

$$B_1 = g_1 \oplus \delta \tag{11}$$

$$B_2 = g_2 \oplus \phi \tag{12}$$

Stage 4: Support value is computed by using the calculated helper values for the encryption which is shown in Eq. (13)

$$S = (B_1 + B_2)\big/B_1 * B_2 \tag{13}$$

Stage 5: Challenge value is computed as $C_V \leftarrow H\left(g_{\text{pk}}, N, S, B_1, B_2\right)$, where hash function is represented as H.
Stage 6: The signature output is as, $\sigma \leftarrow (\delta, \phi, C_V, B_1, B_2)$.

3.4.3 Verification of Key Signature

In signature verification, the SV based signature verification is used to check the nodes. The SV based signature verification will be described in the following stages.

Stage 1: By using group public key, message, and support value, the signature verification nodes' steps are initialized, i.e., Verify$\left(g_{\text{pk}}, \sigma, N\right)$.

Stage 2: Re-drive the $\hat{B}_1$, $\hat{B}_2$ (helper values) to the first step by Eqs. (14) and (15) in the signature verification.

$$\hat{B}_1 = B_1 \oplus \delta \tag{14}$$

$$\hat{B}_2 = B_2 \oplus \phi \tag{15}$$

Stage 3: By Eq. (16), the support value is computed

$$\hat{S} = \frac{\left(\hat{B}_1 + \hat{B}_2\right)}{\left(\underset{1}{\hat{B}} * \hat{B}_2\right)} \tag{16}$$

Stage 4: By using $\hat{C} \leftarrow H\left(g_{\text{pk}}, N, \hat{S}, \hat{B}_1, \hat{B}_2\right)$, challenge value can be determined. Where hash function is represented as H.

Stage 5: Observe the challenge $\left[C_V == \hat{C}_V\right]$.

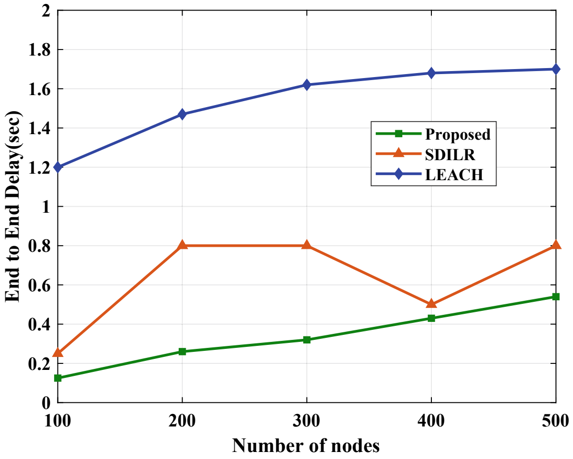
The check step will be valid σ where the challenge value (CV) is the same, and the data is transmitted with the user-verified signature. If the CV is not equal, the user results in an invalid σ, so that the data are not shared with the user and are treated as an attack. By using this, information is safely shared with the verified nodes, and access to the other nodes is not possible. This operation provides the security of information sharing without a tapping attack by the primary user's altogether.

4 Results and Discussion

This work is carried out in MATLAB platform, along with the performance results of PDR, E2E delay, accuracy of detection, energy dissipation, and TP. To show the efficiency of the proposed method, work is compared to existing techniques like improved LEACH routing (SDILR) and low-energy adaptive clustering hierarchy, such as the secure distance-based method (LEACH). Compared to the secure data aggregation (SCDA), Gaussian reputation system (GRSSN), and hierarchical trust management (HTM), the performance of energy dissipation and trust probabilities is comparable.

The E2E performance shown in Fig. 2 is calculated by varying numbers of nodes. Our proposed method is compared with SDILR and LEACH techniques. At the node

Fig. 2 Performance of E2E delay



500, the existing technique of LEACH and SDILR is 1.7 and 0.8 s, and proposed method is 0.5 s. Compared to existing techniques, our proposed work achieves less delay performance.

The throughput performance is calculated by varying number of nodes which is shown in Fig. 3. Our proposed method is compared with SDILR and LEACH techniques. At the node 500, the throughput performance of existing techniques is LEACH and SDILR which is 10,500 bits and 10,800 bits, and proposed method is 11,400 bits. Compared to existing techniques, our proposed work achieves high throughput performance.

The performance of proposed PDR is shown in Fig. 4 by comparing existing techniques like SDILR and LEACH. The PDR is calculated by different number

Fig. 3 Performance of throughput

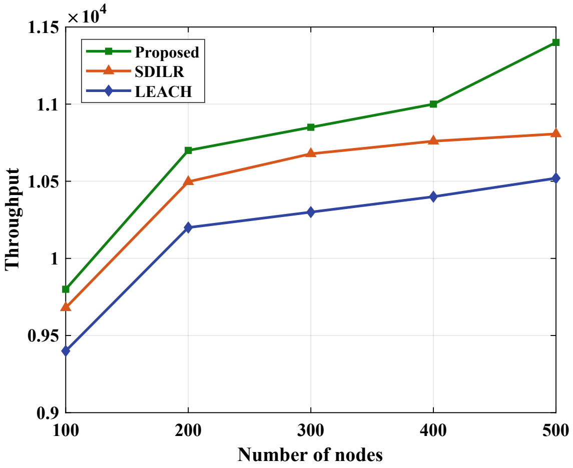
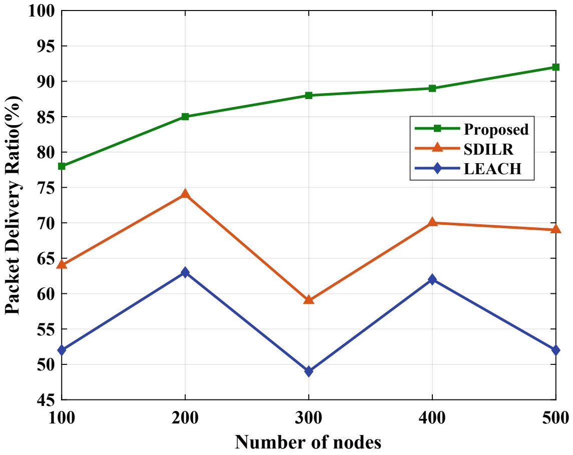


Fig. 4 PDR performance



of nodes. The existing SDILR and LEACH for PDR is 52% and 69.6%, and the proposed technique is 92%. Our proposed method achieves high PDR than existing techniques.

Detection accuracy performance is depicted in Fig. 5 by varying number of nodes. At the node 500, the detection accuracy performance of existing techniques are LEACH and SDILR which is 82 and 93.5%, and proposed method is 97%. Compared to existing techniques, our proposed work achieve high detection accuracy.

The performance of energy dissipated is shown in Fig. 6 by comparing existing techniques like SCDA, GRSSN, and HTM. Different number of MNs is determined by the energy dissipation. The existing SCDA, GRSSN, and HTM for energy dissipation is 34, 41, and 46 J, the proposed technique is 30 J.

Fig. 5 Performance of detection accuracy

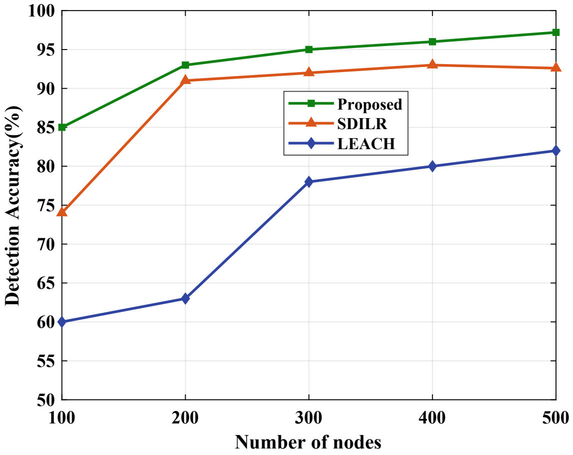
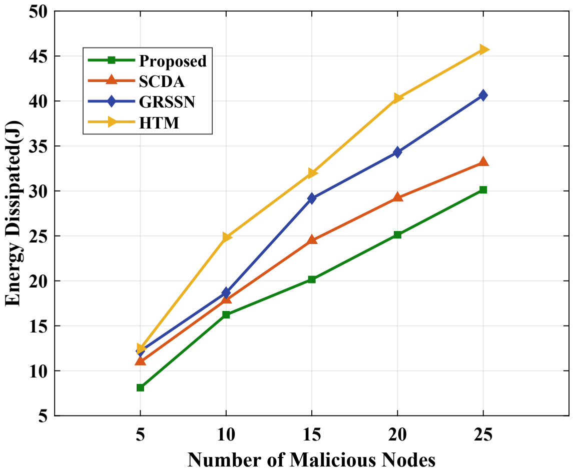
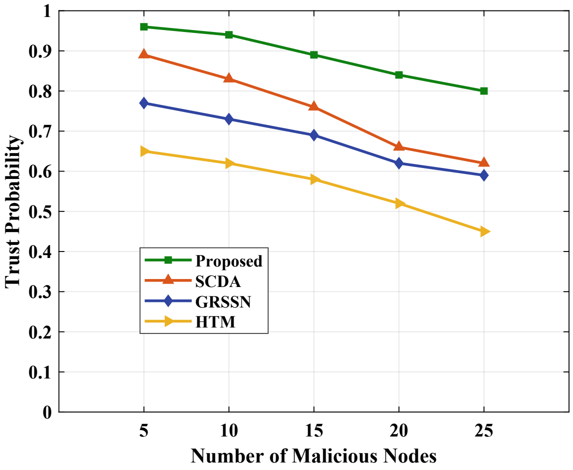


Fig. 6 Energy dissipated performance



The performance of TP is shown in Fig. 7 by comparing existing techniques like SCDA, GRSSN, and HTM. TP is calculated by varying number of MNs. Most of the nodes are dropping its energy soon in SCDA, GRSSN, and HTM. With an increasing number of MNs and fewer regular nodes, the existing method calculates the TP well. Our proposed method calculates the trust properly with proper cluster head with higher amount of regular nodes.

Fig. 7 Number of MNs verses TP



5 Conclusion

In this paper, we addressed the security issues by using STELR protocol. Initially, the sensor node is placed in the network for secure communication without attacking eavesdropping. Here, cluster head is selected by using EPO algorithm based on distance and energy for data aggregation, and STELR is used to enable the security for routing. The work is implemented in MATLAB platform, and the results of the PDR, E2E delay, throughput, detection accuracy, energy dissipated, and TP are achieved. At the node 500, the E2E delay is 1.7 s, throughput is 11,400 bits, PDR is 92% and detection accuracy is 97%. In future, trust model will be extended with meta-heuristic optimization algorithm for reduce the energy consumption.

References

1. W. Fang, X. Wen, J. Xu, J. Zhu, CSDA: a novel cluster-based secure data aggregation scheme for WSNs. Clust. Comput. **22**(3), 5233–5244 (2019)
2. M. Shobana, R. Sabitha, S. Karthik, Cluster-based systematic data aggregation model (CSDAM) for real-time data processing in large-scale WSN. Wireless Pers. Commun. **4**:1–19 (2020)
3. M.R. Senouci, A. Mellouk, A robust uncertainty-aware cluster-based deployment approach for WSNs: coverage, connectivity, and lifespan. J. Network Comput. Appl. **146**, 102414 (2019)
4. H. Fu, Y. Liu, Z. Dong, Y. Wu, A data clustering algorithm for detecting selective forwarding attack in cluster-based wireless sensor networks. Sensors **20**(1), 23 (2020)
5. S. Sujanthi, S.N. Kalyani, SecDL: QoS-aware secure deep learning approach for dynamic cluster-based routing in WSN assisted IoT. Wireless Pers. Commun. **114**(3), 2135–2169 (2020)
6. A.S. Reegan, V. Kabila, Highly secured cluster based WSN using novel FCM and enhanced ECC-ElGamal encryption in IoT. Wireless Pers. Commun. 1–17 (2021)
7. K. Haseeb, A. Almogren, N. Islam, I. Ud Din, Z. Jan, An energy-efficient and secure routing protocol for intrusion avoidance in IoT-based WSN. Energies **12**(21), 4174 (2019)
8. V. Gomathy, N. Padhy, D. Samanta, M. Sivaram, V. Jain, I.S. Amiri, Malicious node detection using heterogeneous cluster based secure routing protocol (HCBS) in wireless adhoc sensor networks. J. Ambient. Intell. Humaniz. Comput. **11**(11), 4995–5001 (2020)
9. P. Sharmila, A.S.M. Priyadharson, A cluster-based secured data transmission protocol for efficient data gathering in WSN. Int. J. Veh. Inf. Commun. Syst. **4**(4), 331–343 (2019)
10. T. Wang, K. Hu, X. Yang, G. Zhang, Y. Wang, A trust enhancement scheme for cluster-based wireless sensor networks. J. Supercomput. **75**(5), 2761–2788 (2019)
11. V. Vijayalakshmi, A. Senthilkumar, USCDRP: unequal secure cluster-based distributed routing protocol for wireless sensor networks. J. Supercomput. **76**(2), 989–1004 (2020)
12. T. Kalidoss, L. Rajasekaran, K. Kanagasabai, G. Sannasi, A. Kannan, QoS aware trust based routing algorithm for wireless sensor networks. Wireless Pers. Commun. **110**(4), 1637–1658 (2020)
13. M. Pavani, P.T. Rao, Adaptive PSO with optimised firefly algorithms for secure cluster-based routing in wireless sensor networks. IET Wireless Sens. Syst. **9**(5), 274–283 (2019)
14. S.S. Ranjani, S. Radhakrishnan, C. Thangaraj, Secure cluster based data aggregation in wireless sensor networks, in *2014 International Conference on Science Engineering and Management Research (ICSEMR)* (IEEE, 2014), pp. 1–6
15. K.V. Kumar, T. Jayasankar, V. Eswaramoorthy, V. Nivedhitha, SDARP: security based data aware routing protocol for ad hoc sensor networks. Int. J. Intell. Networks **1**, 36–42 (2020)

16. B. Sahu, P. Parida, A.K.. Parida, S.K. Mishra, Token based data security in inter cluster communication in wireless sensor network, in *2020 International Conference on Computer Science, Engineering and Applications (ICCSEA)* (IEEE, 2020), pp. 1–6
17. R. Bestak, Intelligent traffic control device model using Ad Hoc network. J. Inf. Technol. **1**(02), 68–76 (2019)
18. A. Chandy, Smart resource usage prediction using cloud computing for massive data processing systems. J. Inf. Technol. **1**(02), 108–118 (2019)