Wireless Personal Communications (2020) 114:813–833
https://doi.org/10.1007/s11277-020-07394-z

An Integrated Trust Assisted Energy Efficient Greedy Data Aggregation for Wireless Sensor Networks

K. P. Uvarajan[1] · **C. Gowri Shankar**[2]

Published online: 24 April 2020

Abstract
Wireless sensor networks (WSN) gathers information pertaining to sensitive data. As because the sensor nodes in WSN are remote and unattended for a longer period of time, vulnerability to intrusions and attacks are also found to be higher, hence making the medium insecure. However, increasing reliability also results in an increase in energy consumption significantly. Though several methods and mechanisms designed for solving the above said security and energy issues, many of these experiences notable computational, communication, and storage requirements that frequently cannot be contended by resource-constrained sensor nodes. Therefore, in this work, a threefold homogeneous method is introduced that supervises secure neighbor selection, energy-efficient routing and data aggregation with the greedy approach that provides multi-objective purpose called, Trust Assisted Global and Greedy Congestion-aware Data Aggregation for (TAG–GCDA) secured WSN with reduced energy consumption and improved reliability. The method enhances global aggregation precision with finite restrictions in neighbor reliability and aggregation. The threefold process of the TAG–GCDA method ensures trusted neighbor selection based on correlative divergence, energy conservation using the global cost (i.e. predictable trust and unpredictable trust) and greedy congestion control for seamless transmission. These processes intend to reduce the energy usage of the sensors to increase the network lifetime with lesser control and communication overhead. The trade-off between energy and security is obtained so as to advance efficient energy consumption with a higher packet delivery ratio.

Keywords Wireless sensor networks · Trust assisted · Global · Greedy congestion-aware · Data aggregation

✉ K. P. Uvarajan
uvaraj.kp@gmail.com

C. Gowri Shankar
cgshankar@gmail.com

[1] Department of ECE, K.S.R. College of Engineering, Namakkal, Tamilnadu, India

[2] Department of EEE, K.S.R. College of Engineering, Namakkal, Tamilnadu, India

1 Introduction

WSNs possess self-dependent sensors scattered in the entire network that are straightforwardly adaptable in adverse situations to monitor environmental conditions like, noise, temperature, pressure, humidity and so on. Due to the existence of a large number of sensors in WSN, security remains one of the major aspects to be covered.

One of the most prevailing solutions for ensuring security in WSN is the application of cryptography. By applying cryptography, the key is distributed between the sensors either in a symmetric or asymmetric pattern. With the objective of minimizing the complexity and remove the key distribution between sensors, a Hamming residue method (HRM) was presented in [1]. The HRM increased the security of network against malicious attacks and also boosts the network efficiency. Initially, a codeword was said to be generated with the help of defined initial security bits and security check bits via Hamming bits. Followed by which at several hops, the quadratic residue was applied to increase the security.

Finally, with the aid of IPV6, the information pertaining to HRM was stockpiled in the header and upon successful code matching, the data was said to be accessed or else the sensor was considered as a malicious node. With this, the security of WSN was said to be enhanced by applying the Hamming residue technique. Since codeword was generated at each node, therefore confidentiality was also said to be improved, therefore increasing the packet delivery ratio and minimizing the delay. Though security was said to be ensured, however, the energy consumed during confidential packet delivery remained unaddressed.

With the inception of WSNs in recent years, administering trustworthy and reliable data delivery is a demanding function due to distinctive features and sensor limitations. An efficient belief-based trust evaluation mechanism (BTEM) was presented in [2] that separated the malicious sensor from trust-worthy sensors. Initially, the Bayesian estimation mechanism was applied for obtaining direct and in-direct trust factors of sensors. Besides, it also scrutinizes the data correlated over time. In this manner, the mechanism estimated coarse-grained knowledge for the purpose of decision making that ensured secure data delivery and hence circumventing the malicious sensors. With this, the BTEM performed better in terms of detection rate, identifying trustworthiness sensors with lesser delay and improved network throughput. Though secured data delivery was said to be performed, the energy consumed for secured data delivery was not focused.

With the increase in the aging population, a medical information system based on the telecare has become a vital necessity to address issues related to chronic diseases, mental illnesses, cost and stress of the frequent referral to the hospital, and so on. An IP-based secured routing protocol was designed in [3] using an adapted distribution scheme. The scheme was found to be efficient in terms of energy being consumed and also improving security using compression and distribution models. Yet another method to be attended for medical healthcare was designed in [4] by applying the elliptic curve via a self-certified key management scheme. With this, security was said to be improved.

One of the crucial roles in several domains is the detection of anomaly or malicious nodes. In [5], a one-class support vector machine was applied for anomaly detection. Besides, a random approximation function was also used in addition to stochastic gradient descent to reduce both the running time and memory consumption involved in anomaly detection. However, the source location privacy was not focused on. To address this issue, source location privacy was concealed in [6] using multiple sinks. With this, besides security, network lifetime along with the energy consumption was also found to be reduced. A survey of data collection methods for measuring security in WSN was presented in [7].

Yet another method for protecting source location was designed in [8] by applying sector-based random routing (SRR). By applying SRR, data packets were transmitted to random phantom sources located in numerous sectors to reach the sink node via different directions. Besides, to minimize the consumption of energy, the hop threshold was also used. However, security was less focused. With the solution being a searchable public-key encryption method, in [9], a Lightweight Searchable Public-key Encryption (LSPE) scheme was presented to reduce considerably both the time and cost involved in performing encryption. In WSN, data aggregation permits in-network processing, which in turn results in minimizing the data packet transmission rate and also minimizes redundancy. Several research works have been conducted using the Elliptic Curve ElGamal homomorphic encryption algorithm to protect data confidentiality. In [10], Okamoto–Uchiyama homomorphic encryption algorithm was used to safeguard the data confidentiality besides minimizing energy consumption.

In this work, we intend to provide secure WSN performance using an integrated threefold homogeneous method, stable neighbor selection using Trustworthiness Reinforced Neighbor Selection (TRNS) model, energy-efficient routing using Global Cost-based Energy Conservation (GCEC) model and congestion-aware data aggregation using Greedy Congestion-aware Data Aggregation (GCDA) model. Besides the computational and communication overhead caused due to integrating the independent models is intended to be minimized. The remainder of the manuscript is systematized as follows. Section 2 describes the review of the related works. Section 3 describes the network model, problem definition and the proposed method with the threefold homogeneous process in detail. Section 4 deals with the simulation results and discussion. Finally, the conclusion is briefed in Sect. 5.

2 Related Works

In WSN, the existing data aggregation scheme specifically cannot measure the probability of data damage. To alleviate those issues, a resilient data aggregation based on spatiotemporal correlation was designed in [11]. Based on the distributed data convergence method, the resilient data aggregation method integrated the centroid distance and similarity with the objective of evaluating the degree of attack to increase the precision of data recovery being made. However, the security aspects were not covered. An end-to-end secure key distribution model was designed in [12] to not only improve security but also with acceptable overhead. However, the existence of attacks at different layers was not addressed. To provide a solution to this, a Layer Trust-based Intrusion Detection System was presented in [13] with the objective of improving the detection accuracy and reduce the false positive rate concurrently.

Inspired by military applications like battle surveillance, the growth of WSN, in the present days, networks are employed in several industrial and consumer applications, to name a few are, monitoring and control of the process involved in industrial design, monitoring of health using machines and so on. In [14], secure distribution of patient data was performed by applying Paillier and Elgamal cryptosystems. Besides statistics data using the patient information was also performed without affecting privacy. Yet another authentication scheme based on temporal credential was presented in [15] by applying the session key between the user and sensor nodes. With this, not only security was said to be improved but also the communication overhead was reduced in a significant manner. However, issues

related to routing were not focused. To address this issue, in [16], the Dirichlet distribution function was used to select the next-hop of routing. With this, the method possessed the advantages of both packet delivery ratio and network lifetime.

The pursuit to discern real-world occurrence at a fine spatial–temporal resolution has resulted in a great increase in WSNs. A smart card-based user authentication scheme was presented in [17] to achieve user anonymity. Besides, the scheme also used elliptic curve cryptography to anonymous user-to-gateway authentication. In this way, it allowed sensors to perform lightweight cryptographic operations. A survey to deploy a complete refinement in all operative phases of a WSN like, positioning of sensors, coverage of the network, sensor node clustering and data aggregation was presented in [18] based on genetic algorithm. A survey of security and privacy in WSN was elaborated in [19] for healthcare applications. Wireless technology has significantly improved due to its cost-effectiveness compared to wired scenarios, specifically considering the uses provided by WSN based applications. Such WSN based applications are said to be found in several areas mainly, healthcare. In [20], a health monitoring system was analyzed with respect to power consumption and security aspects.

Different from the conventional method of regulating security in WSNs, our contribution is formulated as below:

- To increase the network lifetime, the proposed TAG–GCDA method utilizes the TRNS model. Trustworthiness Reinforced Neighbor Selection (TRNS) model in which a neighbor sensor is obtained for its cohesive time with the trustworthiness value. This, in turn, identifies specific neighbors with lesser communication overhead that in turn perform reliable data packet communication over a prolonged time.
- To minimize the energy consumption, the proposed TAG–GCDA method with help of GCEC model. Global Cost-based Energy Conservation (GCEC) model for reducing energy consumption by changing the conventional trustworthiness through its packet forwarding. In the GCEC model, a global cost is evaluated for identifying the route path based on their gained trust value.
- To increase the packet delivery ratio, the GCDA model is introduced in the proposed TAG–GCDA method. Greedy Congestion-aware Data Aggregation (GCDA) model aids in congestion-free data relocation between nodes to enhance reliability in communication. It accounts for the congestion score and probability of packet drop rate to ensure seamless transmissions, improving the packet delivery ratio.

3 Trust Assisted Global and Greedy Congestion-Aware Data Aggregation for Secured WSN

The applications of WSNs increase from military, health, area monitoring, forest fire detection and so on. Hence, WSNs are utilized in monitoring and tracking elders and patients for health care purposes, surveillance that notably alleviate the critical inadequacy of health care personnel, target tracking and minimize the health care expenses along with addressing security in the present-day health care systems.

For example, to monitor the patient's behavior, sensors are deployed in a patient's home whereas an alert message is sent to the doctors when the patient falls ill and requires immediate medical care. On the other hand in surveillance, critical events are also recorded. In this section, the proposed method, Trust Assisted Global and Greedy Congestion-aware

Data Aggregation for (TAG–GCDA) secured WSN is designed for healthcare in elaborate followed by the network model and problem definition. Figure 1 shows the block diagram of the TAG–GCDA method.

As shown in the above figure, the Trust Assisted Global and Greedy Congestion-aware Data Aggregation for (TAG–GCDA) secured WSN and enhanced reliability with reduced energy consumption. In the proposed method, three different models are designed such as Trustworthiness Reinforced Neighbor Selection (TRNS) model for increasing network lifetime, Global Cost-based Energy Conservation (GCEC) model for reducing energy consumption and Greedy Congestion-aware Data Aggregation (GCDA) model for increasing packet delivery ratio. The elaborate description of the three different models is given below.

3.1 Network Model

In this work, a distributed network [21] comprising of a large number of sensors that are deployed in a uniform and random manner. The network is modeled as a connected graph '$G = (V, E)$', with 'V' representing the vertices and 'E' representing the links, each vertex corresponding to a network node and each link corresponding to a communication channel. Here, each sensor is limited by computing power and energy. There always exists a trusted node '$SN \in V$' that possess powerful computing and storage capability, with 'SN' known as the sink node. The other nodes are either referred to as the reliable node or unreliable node with the network size being 'N'. Let us further presume that the locations of the sensors remain steady after they have been deployed and sensor node communication is performed via multi-hop routing. In addition, a sink node is situated in the center of the deployment area. Finally, the network is supposed to be an event-monitoring network, with

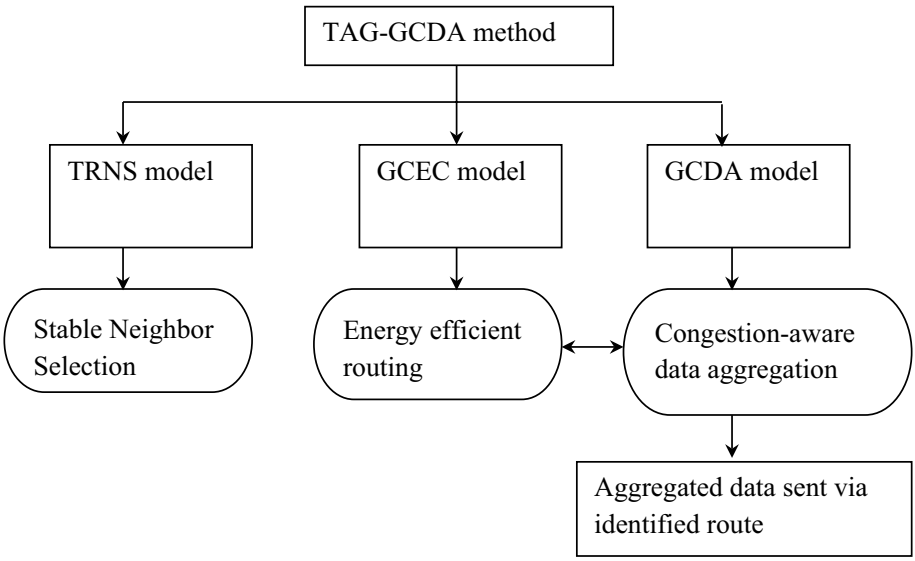

Fig. 1 Block diagram of Trust Assisted Global and Greedy Congestion-aware Data Aggregation for secured WSN

the assumption that whenever an event is detected, the source sensor sends the data gathered to the sink node, with only one source node generated at any time.

3.2 Problem Definition

Handling security against resource constraints in WSN necessitates auxiliary control message for the swift discovery of neighboring sensor nodes (i.e. sensors). Neighbor sensor discovery is persistent due to inappropriate and incorrect neighboring sensor information update due to the inadequate neighbor verification methods. With several data authentication methods in existence, though data transmission is said to be secured, however, data aggregation quality is degraded due to the minimization of transmission rate and accuracy during data gathering. The proposed method reduces the probability of delayed neighbor selection and increases reliability check. The stability of the sensor (i.e. stable sensor node) for estimating its reliability is assessed in a periodical manner. The choice for sensors is given in the order of higher stability. In addition, the trade-off between security and energy is addressed by enhancing the precision involved in data aggregation, using definite processes of data aggregation and energy-efficient routing.

3.3 Trustworthiness Reinforced Neighbor Selection Model

In this section, a Trustworthiness Reinforced Neighbor Selection (TRNS) model is designed with the objective of selecting the neighbor sensor node in a cohesive manner. Figure 2 shows the flow diagram of the TRNS model.

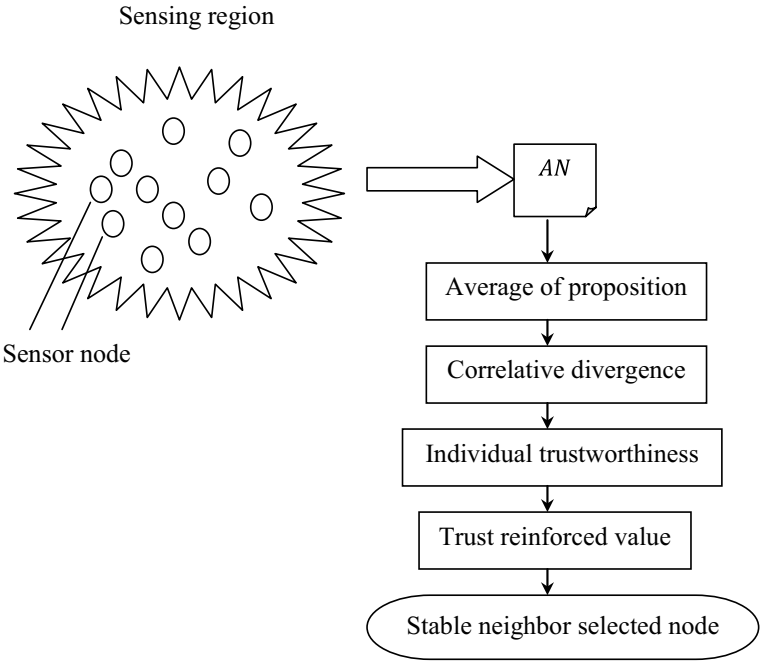

Fig. 2 Flow diagram of TRNS model

Energy consumption remains the major trust metric with which the trustworthiness of any node is measured. Therefore, 'EC' and 'DP' are considered as the trust metrics for trustworthiness in the proposed work. Here, the aggregator node 'AN' collects propositions from its neighbor sensor nodes after time interval 'Δt'. Then, the average of the propositions is measured as given below.

$$P_{EC} = \frac{1}{n} \sum_{i=1}^{n} EC_n \tag{1}$$

$$P_{DP} = \frac{1}{n} \sum_{i=1}^{n} DP_n. \tag{2}$$

From the above Eqs. (1) and (2), 'P_{EC}' and 'P_{DP}' are the average of the propositions collected after time interval 'Δt', with 'n' representing the number of neighboring sensor nodes. The 'EC' provided by the neighboring sensor node to the aggregator node 'AN' is measured by the energy consumed in transmission during the time interval 'Δt' and is represented as given below.

$$EC_{bit} = C * V * t. \tag{3}$$

From the above Eq. (3), 'EC_{bit}', refers to the product of the current 'C', voltage 'V' and time 't', respectively. After calculation of 'P_{EC}' and 'P_{DP}', the correlative divergence of the trust metrics of node 'i' is represented as given below.

$$CD_{EC} = \frac{\Delta EC_i(t) - \Delta P_{EC}(t)}{\Delta P_{EC}(t)} \tag{4}$$

$$CD_{DP} = \frac{\Delta DP_i(t) - \Delta P_{DP}(t)}{\Delta P_{DP}(t)}. \tag{5}$$

From the above Eqs. (4) and (5), 'CD_{EC}' and 'CD_{DP}' corresponds to the divergences of trust metrics.

$$\Delta EC_i(t) = EC_i(t - \Delta t) - EC_i(t) \tag{6}$$

$$\Delta DP_i(t) = DP_i(t - \Delta t) - DP_i(t). \tag{7}$$

From the above Eqs. (6) and (7), '$EC_i(t)$' corresponds to the energy consumed at the time 't' and '$\Delta EC_i(t)$' corresponds to the energy consumed between node 'i' and node 'j' respectively. Now the individual trustworthiness using 'EC_i' and 'DP_i' is measured as given below.

$$T_{ij}^{EC}(t) = \begin{cases} 0, & \text{if } \Delta EC_i(t) > \Delta P_{EC}(t) \\ 1, & \text{Otherwise} \end{cases} \tag{8}$$

$$T_{ij}^{DP}(t) = \begin{cases} 0, & \text{if } \Delta DP_i(t) > \Delta P_{DP}(t) \\ 1, & \text{Otherwise} \end{cases}. \tag{9}$$

From the above Eqs. (8) and (9), it is observed that if '$\Delta EC_i(t)$' and '$\Delta DP_i(t)$' is greater than the average values, the trustworthiness of the node reduces. The final trust reinforced value is measured as given below.

$$T_{ij}(t) = T_{ij}^{EC}(t) + T_{ij}^{DP}(t). \tag{10}$$

The pseudo-code representation of the Average Propositioned Correlative Divergence Neighbor Selection is given below.

Input: Source node '$S = S_1, S_2, \ldots, S_n$', Aggregator node '$AN$', Sink node '$SN$', Energy consumed '$EC$', Number of data packets received 'DP',
Output: Trustworthiness Reinforced node 'TS_i'
1: **Begin**
2: **For** each Source node 'S' with Aggregator node 'AN' and Sink node 'SN'
3: Let energy consumed be 'EC' and number of data packets received 'DP'
4: Measure the average of the propositions using (1) and (2)
5: Measure correlative divergence of the trust metrics of node 'i' using (4) and (5)
6: Measure individual trustworthiness using (8) and (9)
7: Measure the final trust reinforced value using (10)
8: Return (high trustworthiness reinforced node)
9: **End for**
10: **End**

Algorithm 1 Average Propositioned Correlative Divergence Neighbor Selection

As given in the above Average Propositioned Correlative Divergence Neighbor Selection algorithm, there exist three different nodes, called, source node, aggregator node and sink node for the selection of stable neighbor selection. Here, first, the average proposition between the aggregator and the neighbor sensor nodes is measured. Followed by which the correlative divergence between the aggregator and the neighbor sensor nodes is obtained. With the average proposition and correlative divergence values, individual trustworthiness and the final reinforced trust values are measured with which the neighbor sensor node is selected. In this way, the proposed method minimizes the probability of delayed neighbor node selection and enhances the reliability of the network. Upon successful selection of the neighbor sensor node, the source node relays all of its available data packets to the aggregator node via the neighboring sensor node. This selected neighbor sensor node is considered a trusted node. The trust of the neighbor sensor node is refreshed upon transmission and new trustworthiness value is obtained.

3.4 Global Cost-Based Energy Conservation Model

With the identified Trustworthiness Reinforced node, the next objective in the proposed method is to reduce the energy being consumed using an efficient energy conservation

model. In this work, a Global Cost-based Energy Conservation (GCEC) model is designed. The GCEC model identifies a route from source to a destination node via the sink node in a dynamically changing network by applying global cost. In order to identify a secure and energy-efficient route, the proposed method alters the route request frame '*RREQ*' and route reply frame '*RREP*' by adding the information pertaining to both trust and residual energy.

The information pertaining to trust and residual energy is exchanged without enlarging the communication traffic by applying the Global Cost-based Energy-efficient Routing (GCER) algorithm. The GCER algorithm defines the route global cost '*RGC*', the source node selects a better forwarding route by comparing the '*RGC*'. route global cost '*RGC*' scrutinizes the trust value, the residual energy (the energy left after completion of the entire routing process of the networks) and the count of hope, and '*RGC*' is measured as given below.

$$W_{RE} = Er \sum_{i=1}^{n} \left[TS_i\right]. \tag{11}$$

From the above Eq. (11), the residual energy '*RE*', is measured based on the remaining energy '*Er*' of sensors 'S_i'. Next, the Global Cost (i.e. predicted trust and unpredicted trusts) are mathematically formulated as given below.

$$PT_{ij} = \frac{\alpha_{ij}}{\alpha_{ij} + \beta_{ij}} \tag{12}$$

$$UPT_{ij} = 1 - \left[\frac{\alpha_{ij}}{\alpha_{ij} + \beta_{ij}}\right]. \tag{13}$$

From the above Eqs. (12) and (13), the predicted trust 'PT_{ij}' and unpredicted trust 'UPT_{ij}' is evaluated based on the predicted interaction 'α_{ij}' and unpredicted interaction 'β_{ij}' between sensor nodes 'i' and 'j' respectively. With the obtained residual energy from Eq. (11), predicted trust from Eq. (12) and unpredicted trust from Eq. (13), the Route Global Cost '*RGC*' is measured as given below.

$$RGC = W_{RE} + PT_{ij} + UPT_{ij} + W_{HC}. \tag{14}$$

From the above Eq. (14), Route Global Cost '*RGC*', is obtained using the weight of residual energy 'W_{RE}', predictable trust 'PT_{ij}', unpredictable trust 'UPT_{ij}' and weight of the hop count 'W_{HC}' between the node 'i' and 'j' respectively. The pseudo-code representation of Global Cost-based Energy-efficient Routing is given below.

Input: Trustworthiness Reinforced node 'TS_i', Route request '$RREQ$', Route reply '$RREP$', Threshold 'Th_E'
Output: Identification of Cost-based Energy-efficient Route
1: **Begin**
2: **For** each Trustworthiness Reinforced node 'TS_i'
3: Broadcast ' $RREQ$' frames and '$RREP$' frames
4: Measure residual energy using (11)
5: Measure predicted trust using (12)
6: Measure unpredicted trust using (13)
7: Measure Global Cost of Route using (14)
8: **End for**
9: **End**

Algorithm 2 Global Cost-based Energy-efficient Routing

As given in the above Global Cost-based Energy-efficient Routing algorithm, the objective remains in identifying the energy-efficient routing based on the trust and residual energy with the objective of providing secure routing. To start with, the GCEC model broadcasts '*RREQ*' frames and '*RREP*' frames during route construction. When a sensor node with high trust and low residual energy acquires '*RREQ*' frame, this sensor node chooses to discard the '*RREQ*' frame. This is because of the reason that its residual energy is lesser than the specified threshold 'Th_E'. When a sensor node with a low trust and a high residual energy reply a '*RREP*' frame to the requested sensor, the requested sensor will discard the '*RREP*' frame to avoid building a route which include includes with low trust. Finally, the route with the lowest '*RGC*' which is considered to be the secured route is chosen to transmit data. By selecting a trustworthy sensor for data transmission, the GCEC model improves security. By selecting the distinctly residual energy of sensor nodes for routing, the GCEC model minimizes the energy consumption of a high trust node and reduces the route break caused by insufficient energy to improve routing stability.

3.5 Greedy Congestion-Aware Data Aggregation Model

Finally, with the identified trustworthy route, the last model designs the data aggregation using the greedy approach and the aggregated data is sent to the destined node via the identified trustworthy route. The model is called the Greedy Congestion-aware Data Aggregation (GCDC) model. Besides, we said that the method is greedy as it considers both the congestion score and probability of data packet being dropped. This is performed using the Congestion Score '*CS*' for each trust reinforced node 'TS_i' and is mathematically formulated as given below.

$$CS(TS_i) = \left[\frac{TDP[IN]TS_i}{DP[IN]TS_i}\right]t. \tag{15}$$

From the above Eq. (15), the congestion score is obtained based on the ratio of total data packet input score '$TDP[IN]TS_i$' to the data packet forwarding rate '$DP[IN]TS_i$' for each trust reinforced node 'TS_i' for time interval 't'.

$$TDP[IN]TS_i = S_{source} + S_{intermediate}. \tag{16}$$

From the above Eq. (16), the data packet input score '$TDP[IN]$' refers to the summation score of corresponding source traffic 'S_{source}' and intermediate traffic '$S_{intermediate}$' by trust reinforced node 'TS_i' for time interval 't'. Besides, obtaining the congestion score, the proposed method also evaluates the probability of the data packet drop in a greedy manner. The probability of data packet drop rate is mathematically evaluated for two different conditions and is expressed as given below.

$$Prob_d = \frac{\rho^{l+1} - \rho^l}{\rho^{l+1} - 1}, \quad if\ \rho \neq 1 \tag{17}$$

$$Prob_d = \frac{1}{l+1}, \quad if\ \rho = 1. \tag{18}$$

From the above Eq. (17) and (18), with '$\rho = \lambda/\mu$' being the ratio of data packet arrival rate 'λ' and data packet service rate 'μ', the probability of data packet drop rate '$Prob_d$' is measured for two different conditions with respect to the queue length 'l' (i.e. data packets to be ready for aggregation). The pseudo-code representation of Greedy Congestion-aware Data Aggregation is given below.

Input: total data packet input score '$TDP[IN]TS_i$', data packet forwarding rate '$DP[IN]TS_i$'
Output: improved packet delivery ratio
1: **Begin**
2: **For** each Cost-based Energy-efficient Route
3: Given total data packet input score '$TDP[IN]TS_i$' and data packet forwarding rate '$DP[IN]TS_i$'
4: Measure Congestion Score using (15)
5: Measure Probability of data packet drop score using (17) and (18)
6: **If** '$CS(TS_i) = 0$' and '$if\ \rho\ = 1$'
7: Then congestion happens
8: No data aggregation
9: **End if**
10: **If** '$CS(TS_i) \neq 0$' and '$if\ \rho\ \neq 1$'
11: Then congestion is said to occur
12: Perform data aggregation
13: Aggregated data are transmitted via identified route
14: **End if**
15: **End for**
16: **End**

Algorithm 3 Greedy Congestion-aware Data Aggregation

As given in the above algorithm, first, the algorithm is said to be greedy because it considers both the congestion score and probability of data packet drop score before data aggregation. Next, only by checking the above two values, data aggregation is said to take place. Hence, only if congestion is said to be free, the aggregated data are transmitted via the identified route, or else it is assumed that congestion is said to exist and data are aggregated and only upon congestion-free, the aggregated data is transmitted via identified route. In this way, the packet delivery ratio is said to be improved.

4 Simulation Settings

In this section, the performance of the proposed Trust Assisted Global and Greedy Congestion-aware Data Aggregation for (TAG–GCDA) secured WSN through simulations in comparison with previous methods, Hamming residue method (HRM) [1] and belief-based trust evaluation mechanism (BTEM) [2]. In specific, we measured the communication overhead, energy consumption, and packet delivery ratio under different network conditions. Communication overhead is measured based on overhead incurred during communication. Energy consumption is measured according to the energy be consumed during

secured transmission in WSN. The packet delivery ratio is measured by the percentage of nodes whose sensor readings are received by the sink. The NS2 simulator was used to validate performance evaluation related to the proposed method under consideration. From the experimental results analysis, the proposed TAG–GCDA method achieves secure communication in WSN by evaluating trustworthiness with reduced communication overhead and energy consumption and increased packet delivery ratio. Simulation parameters are provided in Table 1. The results from the simulation show that the effectiveness and the efficiency are tangible with respect to the performance of the proposed method when considering the different parameters.

5 Discussion

We implemented the above four methods in the simulation, which operate on the same aggregation scheme in order to perform comparable experiments. The simulation constructs a data aggregation scheme for secured WSN for each node. The measurement results for each method are the average values over 10 runs.

5.1 Scenario 1: Performance Analysis of Communication Overhead

Communication overhead refers to the total number of packets transmitted from one sensor node to another sensor node. It includes the overhead of the routing process, routing table, and packet preparation. In our work, the communication overhead refers to the memory consumed for Global Cost of Route '*RGC*' and memory consumed for obtaining residual energy. It is mathematically formulated as given below.

$$CO = MEM\,[RGC] + MEM\left[W_{RE}\right]. \tag{19}$$

From the above Eq. (19), communication overhead '*CO*' refers to the memory consumed for residual energy 'W_{RE}' and global cost '*RGC*' respectively. It is measured in terms of kilobytes (KB). Lower the communication overhead, the higher the rate of packet delivery ratio is. The sample calculations for obtaining communication overhead using the proposed TAG–GCDA method, and existing methods HRM and BTEM is given below.

Sample calculation

Table 1 Simulation parameters

Parameters	Values
Simulator	NS-2 (v2.34)
Simulation landscape	1500 m × 1500 m
Number of nodes	500
Transmission range	250 m
Node initial energy	100 J
Packet size	1000 bits
Mobility models communication	Random waypoint
Mobility time	0–30 m/s
Routing protocol	AODV
Packet interval	0.01 s

- *Proposed TAG–GCDA*: The memory consumed for residual energy 'W_{RE}' being '7 KB' and memory consumed for obtaining global cost '*RGC*' being '12 KB', the overall communication overhead is given below.

$$CO = 12\,\text{KB} + 7\,\text{KB} = 19\,\text{KB}$$

- *Existing HRM* The memory consumed for residual energy 'W_{RE}' being '11 KB' and memory consumed for obtaining global cost '*RGC*' being '15 KB', the overall communication overhead is given below.

$$CO = 15\,\text{KB} + 11\,\text{KB} = 26\,\text{KB}$$

- *Existing BTEM* The memory consumed for residual energy 'W_{RE}' being '12 KB' and memory consumed for obtaining global cost '*RGC*' being '19 KB', the overall communication overhead is given below.

$$CO = 19\,\text{KB} + 12\,\text{KB} = 31\,\text{KB}$$

The performance graph of computational overhead using the three methods, Trust Assisted Global and Greedy Congestion-aware Data Aggregation for (TAG–GCDA), Hamming residue method (HRM) [1] and belief-based trust evaluation mechanism (BTEM) [2] are illustrated below. Figure 3 given below shows the convergence graph of communication overhead with respect to 500 numbers of sensor nodes for 150 different packets varying in the range of 50–500 KB.

The convergence graph of communication overhead is shown in Fig. 3. For a 50 sensor node network initially, a communication overhead of 19 KB is used and the communication overhead improves slightly when the sensor node increase from 50 to 100. The communication overhead decreases another time from 26 to 31% when 100 sensor nodes are present in the network. When the network size is increased from 100 to 150, the communication overhead also increases from 31 to 43%. This depicts the performance of the proposed TAG–GCDA method. A decrease is observed in the communication overhead value of the curve from its 19% value with 250 sensor nodes. The better output is produced when the proposed methods make use of the Trustworthiness Reinforced Neighbor

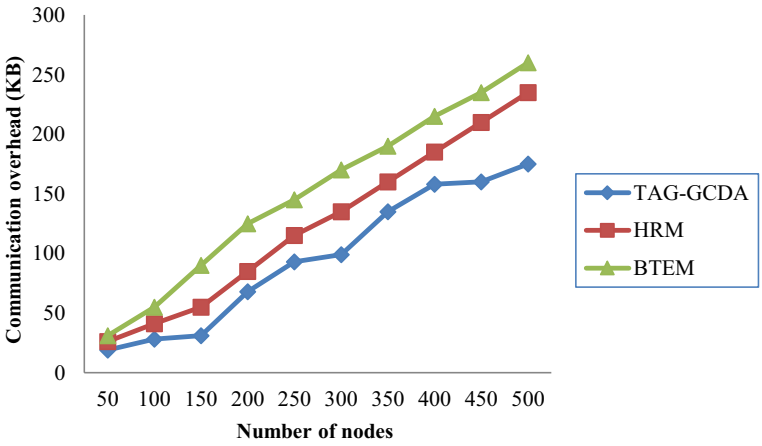

Fig. 3 Performance graph of communication overhead

Selection (TRNS) model in comparison with the existing methods. By applying the TRNS model, propositions from the neighbor sensor nodes with respect to the source node were collected. Besides, the correlative divergence of the trust metrics was also obtained from the neighbor sensor nodes with respect to the source node. With this, individual trustworthiness between nodes and final trust reinforced value was obtained using the Average Propositioned Correlative Divergence Neighbor Selection algorithm. This resulted in the improvement of communication overhead using the proposed TAG–GCDA method by 25% compared to the Hamming residue method (HRM) [1] and 40% compared to the belief-based trust evaluation mechanism (BTEM) [2] respectively.

5.2 Scenario 2: Performance Analysis of Energy Consumption

Energy consumption is measured as the amount of energy that is spent by network nodes within the simulation time. This is obtained by calculating each node's energy level at the end of simulation and factoring in the initial energy of each one. The following formula will produce the value for energy consumption:

$$EC = \sum_{i=1}^{n} (ene(i) - ini(i)). \tag{20}$$

From the above Eq. (20), energy consumption 'EC' is measured based on the difference between the energy at the end of simulation '$ene(i)$' and energy at initial of simulation '$ini(i)$' respectively, with 'i' representing the number of nodes. It is measured in terms of joules 'J'. The sample calculations for obtaining energy consumption using the proposed TAG–GCDA method, and existing methods HRM and BTEM is given below.

Sample calculation

- *Proposed TAG–GCDA* With '50' sensor nodes considered for simulation, the energy at the end of the simulation being '120 J' and energy at the initial of the simulation being '100 J', the energy consumption is measured as given below.

$$EC = 120\,\text{J} - 100\,\text{J} = 20\,\text{J}$$

- *Existing HRM* With '50' sensor nodes considered for simulation, the energy at the end of the simulation being '135 J' and energy at the initial of the simulation being '100 J', the energy consumption is measured as given below.

$$EC = 135\ \text{J} - 100\ \text{J} = 35\ \text{J}$$

- *Existing BTEM* With '50' sensor nodes considered for simulation, the energy at the end of the simulation being '150 J' and energy at the initial of the simulation being '100 J', the energy consumption is measured as given below.

$$EC = 150\ \text{J} - 100\ \text{J} = 50\ \text{J}$$

The performance graph of energy consumption using the three methods, proposed (TAG–GCDA), and existing HRM [1] and BTEM [2] is shown below. Figure 4 given below shows the convergence graph of energy consumption for 500 different numbers of sensor nodes.

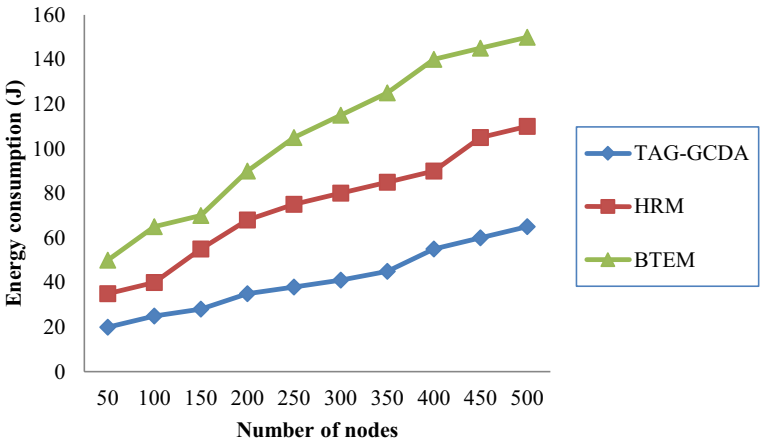

Fig. 4 Performance graph of energy consumption

Figure 4 given above illustrates the convergence graph of energy consumption. Energy consumption here refers to the amount of energy available for communication and data processing. In the network, variances in energy consumption from 50 to 500 sensor nodes are shown in Fig. 4. With the presence of 50 sensor nodes present in the network, the network energy consumption attains a maximum limit of 20 J. When the nodes are increased from 50 to 100 the value of the energy consumption goes proportionally upwards. When the nodes are further increased from 50 to 150, a progressive increase in the maximum energy consumption of the network is observed. However, with the existing HRM and BTEM, the energy consumed for 50 sensor nodes was found to be 35 J and 50 J respectively. Beyond 500 sensor nodes, the maximum energy consumed in the network is 65 J using the TAG–GCDA method, 110 J and 150 J consumed using HRM [1] and BTEM [2]. With this, it is being found that only a relatively minimal value was being observed when applied with the TAG–GCDA method. The average energy consumption using the TAG–GCDA method was found to be comparatively lesser in contrast with a number of other methods including HRM [1] and BTEM [2] methods. This is because of the reason that the TAG–GCDA method employs the Global Cost-based Energy Conservation (GCEC) model that produces superior outputs when compared to other existing methods. By analyzing the Global Cost of Route '*RGC*' (using both predicted trust and unpredicted trusts) along with the residual energy and hop count, the significant route is said to be identified. With the identification of a significant route, transmission flows through this route. This is turn reduces the energy consumption for efficient routing using the proposed TAG–GCDA method by 45% compared to HRM [1] and 61% compared to BTEM [2] respectively.

5.3 Scenario 3: Performance Analysis of Packet Delivery Ratio

Packet delivery ratio refers to the percentage ratio of the data packets that were delivered to the destination node to the data packets that were generated by the source. This metric shows a routing quality in its delivery of data packets from source to destination. Higher the ratio, the better the performance of the method is.

$$PDR = \frac{DP_r}{DP_s} * 100. \tag{21}$$

From the above Eq. (21), the packet delivery ratio '*PDR*', is measured based on the data packets received 'DP_r' and the data packets sent 'DP_s'. It is measured in terms of percentage (%). The sample calculations for measuring packet delivery ratio of three methods, TAG–GCDA, HRM, and BTEM are given below.

Sample calculation

- *Proposed TAG–GCDA* With '15' number of data packets being sent and '13' number of data packets received, the packet delivery ratio is measured as given below.

$$PDR = \frac{13}{15} * 100 = 86.66\%$$

- *Existing HRM* With '15' number of data packets being sent and '12' number of data packets received, the packet delivery ratio is measured as given below.

$$PDR = \frac{12}{15} * 100 = 80\%$$

- *Existing BTEM* With '15' number of data packets being sent and '11' number of data packets received, the packet delivery ratio is measured as given below.

$$PDR = \frac{11}{15} * 100 = 73.33\%$$

Figure 5 given above illustrates the convergence graph of packet delivery ratio obtained at different time intervals. The packet delivery ratio refers to the number of data packets that a sensor node is capable of transferring at a specific time interval. As illustrated in the figure, in the proposed method which has an 86.66% packet delivery when there are 50 sensor nodes (with 15 data packets) in a network and the value of the curve lowers to 83.25% for 100 sensor nodes (with 30 data packets) in the network. When there are a small

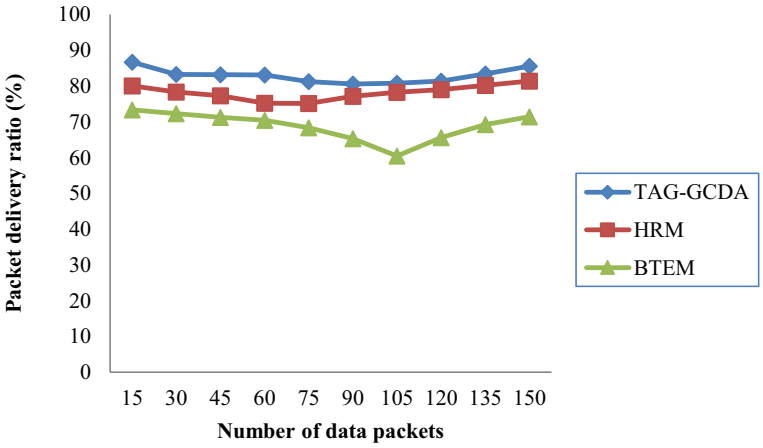

Fig. 5 Performance graph of packet delivery ratio

number of sensor nodes and data packets in a network, the packet delivery ratio is high. Once again, the curve rises to a point where the packet delivery ratio is measured at 85.55% for 500 sensor nodes present in the network. When there are 500 nodes in the network, the curve reaches the value of 85.55%. A greater value of the packet delivery ratio indicates a better performance of the method. The comparison shows betterment using the proposed TAG–GCDA method than the existing HRM [1] and BTEM [2] methods. This is because the TAG–GCDA method used Greedy Congestion-aware Data Aggregation (GCDC), model. By applying the GCDC model, besides measuring the congestion score, the probability of data packet drop rate was also measured using the Greedy Congestion-aware Data Aggregation algorithm. With this, congestion was said to be measured in a greedy manner by considering both the congestion score and probability of data packet being dropped. This, in turn, identifies the congestion if any in an efficient manner and avoids congested routes for the aggregated packets. This, in turn, improves the packet delivery ratio using the TAG–GCDA method by 6% compared to HRM [1] and 21% compared to BTEM [2].

5.4 Scenario 3: Performance Analysis of Packet Loss Rate

Packet LR is measured as the ratio of the number of data packets lost to the total number of data packets sent. PLR is formalized as below,

$$PLR = \left(\frac{DP_L}{N}\right) * 100 \tag{22}$$

From Eq. (22), N denotes the total number of data packets, DP_L represents the number of data packets lost at the designation ends. Packet loss rate measured in Percentage (%).

Sample calculation for packet loss rate:

- *Proposed TAG–GCDA* The number of data packets lost at the destination end is 2 and the total number of data packets sent is 15. Then the packet loss rate is mathematically calculated as,

$$\text{PLR} = \left(\frac{2}{15}\right) * 100 = 13\%$$

- *Existing HRM* The number of data packets lost at the destination end is 5 and the total number of data packets sent is 15. Then the packet loss rate is mathematically calculated as,

$$\text{PLR} = \left(\frac{3}{15}\right) * 100 = 20\%$$

- *Existing BTEM* The number of data packets lost at the destination end is 7 and the total number of data packets sent is 15. Then the packet loss rate is mathematically calculated as,

$$\text{PLR} = \left(\frac{4}{15}\right) * 100 = 26.6{-}27\%$$

The simulation process using 15 data packets sent from the source node, 2 data packets are lost using the TAG–GCDA method obtains 13% and the loss rate of the other two

existing methods HRM [1], BTEM [2] are 20% and 27% respectively. From the above get results, it is significant that the PLR using the proposed TAG–GCDA method is minimal than other conventional methods.

Figure 6 given above illustrates the graph of the packet loss rate. A minimum value of the packet loss rate denotes a better performance of the method. The comparison shows betterment using the proposed TAG–GCDA method than the existing HRM [1] and BTEM [2] methods. This is because the TAG–GCDA method used Greedy Congestion-aware Data Aggregation (GCDC), model. With the help of the GCDC model, then the probability of data packet drop rate was measured. With this measured probability, avoids congested route for the aggregated packets. This, in turn, reduces the packet loss rate using the TAG–GCDA method by 31% compared to HRM [1] and 53% compared to BTEM [2].

6 Conclusion

In WSN, it is critical to guarantee packet delivery accuracy of the physical environmental phenomena. Although many congestion control data aggregation schemes have been proposed to reduce congestion in order to increase packet delivery ratio with minimum consumption of energy, they also concurrently compromise the security due to several attacks. Also, none of them can guarantee packet delivery accuracy. In this paper, we formally analyze the impact of congestion control on the security and packet delivery accuracy aspect. Our analysis results demonstrate the two-sided effect of congestion control on the packet delivery accuracy and the trade-off between security and energy consumption and improving packet delivery accuracy via trust-based model. To reduce the communication overhead while including a trust metric we presented a Trust Assisted Global and Greedy Congestion-aware Data Aggregation for (TAG–GCDA) secured WSN with a packet delivery accuracy guarantee. Based on the average proposition and correlative divergence, the TAG–GCDA method reduces communication overhead. It uses the Average Propositioned Correlative Divergence Neighbor Selection algorithm to reduce communication overhead. TAG–GCDA method also reduces the energy consumption with different route requests based on the global cost, which benefits the accurate environmental data packet monitoring, improving security. Moreover, the TAG–GCDA method improves the packet delivery ratio that ensures security based on the trustworthiness of the node, so that the data packets

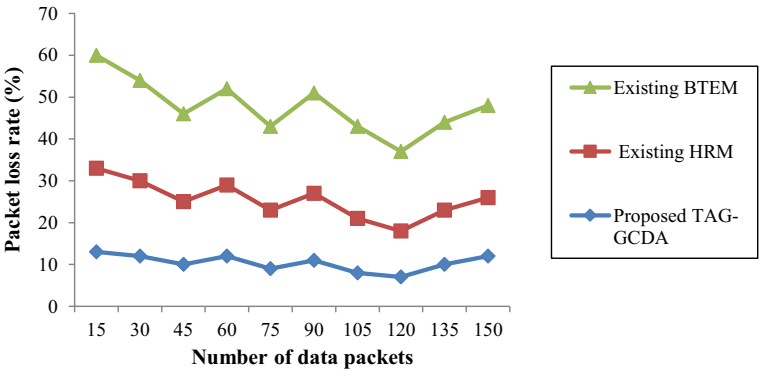

Fig. 6 Performance graph of packet loss rate

delivery is not only ensured but also the trustworthiness of data is ensured. Simulation results prove that the proposed method not only identifies the trustworthy node with which routing is performed but also ensures secure transmission.

References

1. Alotaibi, M. (2019). Security to wireless sensor networks against malicious attacks using Hamming residue method. *EURASIP Journal on Wireless Communications and Networking, 2019*(8), 1–7.
2. Anwar, R. W., Zainal, A., Outay, F., Yasar, A., & Iqbal, S. (2019). BTEM: Belief based trust evaluation mechanism for wireless sensor networks. *Future Generation Computer Systems, 96,* 605–616.
3. Sahraoui, S., & Bilami, A. (2015). Efficient HIP-based approach to ensure lightweight end-to-end security in the internet of things. *Computer Networks, 91,* 26–45.
4. Abbasinezhad-Mood, D., & Nikooghadam, M. (2018). Efficient design of a novel ECC-based public key scheme for medical data protection by utilization of NanoPi Fire. *IEEE Transactions on Reliability, 67*(3), 1328–1339.
5. Miao, X., Liu, Y., Zhao, H., & Li, C. (2019). Distributed online one-class support vector machine for anomaly detection over networks. *IEEE Transactions on Cybernetics, 49*(4), 1475–1488.
6. Han, G., Miao, X., Wang, H., Guizani, M., & Zhang, W. (2019). CPSLP: A cloud-based scheme for protecting source-location privacy in wireless sensor networks using multi-sinks. *IEEE Transactions on Vehicular Technology, 68*(3), 2739–2750.
7. Xie, H., Yan, Z., Yao, Z., & Atiquzzaman, M. (2019). Data collection for security measurement in wireless sensor networks: A survey. *IEEE Internet of Things Journal, 6*(2), 2205–2224.
8. He, Y., Han, G., Wang, H., Ansere, J. A., & Zhang, W. (2019). A sector-based random routing scheme for protecting the source location privacy in WSNs for the Internet of Things. *Future Generation Computer Systems, 96,* 438–448.
9. Xu, P., He, S., Wang, W., Susilo, W., & Jin, H. (2018). Lightweight searchable public-key encryption for cloud-assisted wireless sensor networks. *IEEE Transactions on Industrial Informatics, 14*(8), 3712–3723.
10. Cui, J., Shao, L., Zhong, H., Xu, Y., & Liu, L. (2017). Data aggregation with end-to-end confidentiality and integrity for large-scale wireless sensor networks. *Peer-to-Peer Network Applications, 11*(5), 1022–1037.
11. Yong, L., & Sun, N. (2018). A resilient data aggregation method based on spatio-temporal correlation for wireless sensor networks. *EURASIP Journal on Wireless Communications and Networking, 2018*(157), 1–9.
12. Alves, R. C. A., Oliveira, D. A. G., Pereira, G. C. C. F., Albertini, B. C., & Margi, C. B. (2018). 3: Wireless secure SDN-based communication for sensor networks. *Security and Communication Networks, 2018,* 1–14.
13. Ghugar, U., Pradhan, J., Bhoi, S. K., & Sahoo, R. R. (2019). LB-IDS: Securing wireless sensor network using protocol layer trust-based intrusion detection system. *Journal of Computer Networks and Communications, 2019,* 1–13.
14. Yi, X., Bouguettaya, A., Georgakopoulos, D., Song, A., & Willemson, J. (2015). Privacy protection for wireless medical sensor data. *IEEE Transactions on Dependable and Secure Computing, 13*(3), 369–380.
15. Liu, X., Zhang, R., & Liu, Q. (2017). A temporal credential-based mutual authentication with multiple-password scheme for wireless sensor networks. *PLoS ONE, 36*(1), 316–323. https://doi.org/10.1371/journal.pone.0170657.
16. Zhang, G., Zhang, Y., & Chen, Z. (2013). Using trust to secure geographic and energy aware routing against multiple attacks. *PLoS ONE, 8*(10), 1–7.
17. Nam, J., Choo, K.-K. R., Han, S., Kim, M., Paik, J., & Won, D. (2015). Efficient and anonymous two-factor user authentication in wireless sensor networks: Achieving user anonymity with lightweight sensor computation. *PLoS ONE, 10*(4), 1–22. https://doi.org/10.1371/journal.pone.0116709.
18. Norouzi, A., & Zaim, A. H. (2014). Genetic algorithm application in optimization of wireless sensor networks. *The Scientific World Journal, 6*(4), 152–166.
19. Al-Janabi, S., Al-Shourbaji, I., Shojafar, M., & Shamshirband, S. (2017). Survey of main challenges (security and privacy) in wireless body area networks for healthcare applications. *Egyptian Informatics Journal, 18*(2), 113–122.

20. Dey, N., Ashour, A. S., Shi, F., Fong, S. J., & Sherratt, R. S. (2017). Developing residential wireless sensor networks for ECG healthcare monitoring. *IEEE Transactions on Consumer Electronics, 63*(4), 442–449.
21. Zhang, P., Wang, J., Guo, K., Wu, F., & Min, G. (2017). Multi-functional secure data aggregation schemes for WSNs. *Ad Hoc Networks, 69,* 86–99.

K. P. Uvarajan is an Assistant Professor at K.S.R. College of Engineering, Tamilnadu, India and pursuing Ph.D. degree in Anna University, Chennai. His research interest includes Adhoc networks and embedded systems. He has published many research outcomes in several conferences and peer-reviewed international journals and has about more than 10 research publications to his credit. He is an active member of ISTE and IAENG.

C. Gowri Shankar received his Bachelor's Degree in Electrical and Electronics Engineering from Annai Mathammal Sheela Engineering College, Namakkal, Tamilnadu, India in 2003 and Master's Degree in Applied Electronics from Bannari Amman Institute of Technology, Sathyamangalam, Erode, Tamilnadu, India in 2005 and Ph.D. in the area of Medical Image Processing from Anna University, Chennai in 2013. From 2005 to 2010, he worked as an Assistant Professor in the Department of Electrical and Electronics Engineering, Velalar College of Engineering and Technology, Erode, Tamilnadu, India. Currently, he is working as an Associate Professor in Department of Electrical and Electronics Engineering, K.S.R. College of Engineering, Tiruchengode, Tamilnadu, India, from 2010 to till date. His research interests are Multirate signal processing, Computer vision, Medical image processing and Pattern recognition. He has guided many under graduate, post graduate students and research scholars and he is a Life member of Indian Society for Technical Education (ISTE).