Secured Data Transmission and Malicious Node Detection in Wireless Sensor Network

Deepak N. Biradar, Dr. T.S. Vishanath

Abstract: Wireless Sensor Network (WSN) is developed extremely because of their low installation cost and various applications. WSN has compact and inexpensive sensor nodes for monitoring the physical environment. WSNs are susceptible to many attacks (e.g., malicious nodes) because of its distinct characteristics. The performance of node and network is affected by the malicious nodes. Moreover, the communication among the sensor nodes also required to be secured for preventing the data from the hackers. In this paper, the architecture of the WSN is generated by using the Fuzzy-C-Means clustering (FCM). Then the detection of the malicious nodes is performed by using the Acknowledgement Scheme (AS). This AS is integrated in the Ant Colony Optimization (ACO) based routing for avoiding the malicious nodes while generating the route from the source to the Base Station (BS). Then the Hybrid Encryption Algorithm (HEA) is used for performing the secure data transmission through the network and this proposed method is named as HEA-AS. The performance of the HEA-AS method is evaluated in terms of End to End Delay (EED), network lifetime, throughput, Packet Delivery Ratio (PDR) and Packet Loss Ratio (PLR). The proposed HEA-AS method is compared with the existing method called as CTCM to evaluate the effectiveness of the HEA-AS method.

Index Terms: Wireless sensor networks, acknowledgement scheme, hybrid encryption algorithm, packet delivery ratio, ant colony optimization.

I. INTRODUCTION

WSNs are generally the heterogeneous or homogeneous systems that have numerous tiny sensor nodes and a sink [1]. Typically, the WSN comprises of spatially distributed independent sensor nodes which are utilized for observing the physical or environmental conditions such as vibration, temperature parameters, acoustic parameters and pressure. Then this information is transferred to the primary location (i.e., BS) [2]. The multi-hop communication generates the route among the nodes and then the sensed information is transmitted through the generated transmission path. The characteristics of the WSNs are less memory, limited energy consumption, less bandwidth and limited power supply [3]. The WSNs are used in various real world applications such as home security networks, military, environmental control, habitat monitoring, health monitoring and so on [4]. Nevertheless, the sensors in the WSNs are affected due to the malicious attacks, communication link errors, hardware failure, energy depletion, etc. [5]. The attacks created by malicious nodes are considered as one of the main concern in the WSN. These malicious nodes attack the integrity, availability measures of their neighbors and integrity [6].

There are various types of attacks occurred in the WSNs such as Sybil attack, select forward attack, hello flood attack, sinkhole attack, false routing information attack, wormhole attack etc. [7]. The data transmitted from the source to the destination is susceptible to various security attacks. So, secure routing through the network is established to overcome the issues due to the network layer security attacks [8]. Moreover, the communications among WSNs also need to secure by using Key management for various security features such as confidentiality and authentication [9]. Because the nodes of the WSNs are positioned in an unattended manner as well as the WSNs doesn’t have any centralized management centers for monitoring and handling the risks. Additionally, the malicious actions from the unreliable sensors affect the ability and cooperation of the WSNs [10]. The existing methods used for the secure communication/malicious node detection in WSNs are given as follows: lightweight symmetric key cryptography [11], localized encryption and authentication protocol [12], elliptic curve cryptography [13], deterministic key management scheme [14] and malicious aggregator identification [15]. The major contributions of the proposed HEA-AS method are stated as follows:

- FCM clustering is initialized in the network to decrease the routing overhead and energy consumption through the network.
- The malicious nodes are detected and prevented from the network by using the acknowledgement scheme in the ACO routing. Besides the optimal path from the source to the BS is identified by considering the residual energy and distance between the nodes.
- Here, the security over the network is enhanced by using the hybrid encryption algorithm that combines the dual RSA and MD5 algorithm. The integrity and confidentiality of WSNs are maximized by combining these two algorithms.

II. LITERATURE SURVEY

Kavitha, R.J. and Caroline, B.E [16] presented the Communication and Threshold-based Cryptographic Mechanism (CTCM) to obtain the security in WSNs. The permission access is given to...
the sensors to generate the secret key. After generating the secret key, the key is divided into multiple portions to distribute the keys to multiple sensors. The selfish nodes are identified from the group of sensors to provide secure and reliable data transmission through the network. Then the hierarchical threshold-based secret sharing mechanism is used for preventing the network from intruder’s malicious access. The complexity is increased in the key establishment process by dividing the key into multiple portions.

Stephen, R.K., Sekar, A.C. and Dinakaran, K [17] introduced the Sectional Transmission Analysis (STA) for improving the security and transmission reliability. Initially, the network is divided into several sections and then the routes available for each node are discovered by using the STA. The measure of Reliable Transmission Support (RTS) is used for selecting the optimal route among the network. Further, the measure of Malicious Transmission Support (MTS) is used for mitigating the attacks occurred in the network. The RTS considers various objective functions such as amount of transmissions, energy, amount of neighbors, and the amount of successful transmission. But it fails to consider the distance among the nodes in the RTS.

Zhong, H., Shao, L., Cui, J. and Xu, Y [18] improved the security by using the recoverable data aggregation in the network. This recoverable data aggregation has 5 different phases: 1) setup, 2) private key extraction, 3) encrypt-sign, 4) verify-aggregate-sign, and 5) verify-decrypt. The batch signature verification is used in the BS for verifying the integrity of the encrypted data. After that, the BS performs decryption on the collected cipher text and it provides sensed data. The direct data transmission between the CH to the BS leads to increase the energy consumption through the network.

Mehetre, D.C., Roslin, S.E. and Wagh, S.J [19] presented the two-stage security mechanism and dual assurance scheme for creating the secure and trustable transmission path over the WSN. In the two stage scheme, the node that creates the attacks are identified by using the Detection Packet (DP) and it prevents the sensor from the attacker. Further, this two stage scheme is used for creating the trusted path by considering the minimal threshold value. The dual assurance scheme is split into two different types: 1) selective forwarding based packet validation and 2) ECC based packet security. These methods are used for performing secure data transmission through the WSN. The threshold value used in this work considers only the distance among the nodes.

Saravanaselvan, A. and Paramasivan, B [20] introduced the heavy weight security (HWS) algorithm to provide a secure data aggregation. There are three different steps are processed in HWS algorithm such as: 1) The process of node authentication and verification of data, 2) the public key is created based on the node-location and node-ID to encrypt and decrypt the data packets and 3) the collected data packets are transmitted to the BS. Additionally, the soft computing approach also considered in this work to avoid the malicious nodes in the route generation. The average delay of the WSN is high due to the variation in mobility.

III. HEA-AS METHOD

The proposed HEA-AS method is used to develop malicious node detection and secure communication among the WSN. This technique increases the number of packets delivered to the BS by avoiding the malicious nodes during the data transmission process. The HEA-AS method uses the FCM algorithm for clustering the network. Here, the secure communication over the WSN is provided by using HEA that contains dual RSA and MD5. The ACO routing uses AS scheme for detecting and preventing the network from the effects due to the malicious nodes. The flowchart of HEA-AS method is shown in Fig. 1. The operations performed in HEA-AS method is described below.

A. Clustering using Fuzzy C-Means algorithm

The location of the sensor nodes is given as the input to the FCM algorithm to cluster the network. Similar to the K-means algorithm, the FCM also has the objective of cluster division over the network. Generally, the FCM is a centralized clustering algorithm and here the BS calculates and assigns the sensors in the respective clusters. The sensors are assigned in the cluster by considering the position of the sensors and the CH is allocated by the node which has a high amount of residual energy. Consider, there is $M$ amount of sensor nodes which are split into $c$ clusters like $c_1, c_2, c_3, \ldots, c_m$. The objective of the FCM algorithm is to minimize the following Eq. (1).

$$K_s = \sum_{j=1}^{M} \sum_{i=1}^{M} \mu_{ji}^p d_{ji}^2$$

(1)
Where, the \( k^{th} \) node degree for cluster \( j \) is represented as \( \mu_{jk} \) and the distance among the \( k^{th} \) node and cluster center is denoted as \( d_{jk} \).

The following Eq. (2) is used to identify the cluster centroid of the \( k^{th} \) cluster.

\[
x_k = \frac{\sum_{j=1}^{M} \mu_{jk} d_j}{\sum_{j=1}^{M} \mu_{jk}}
\]  

(2)

The Eq. (3) is used for calculating and fuzzifying the \( k^{th} \) node degree \( (\mu_{jk}) \). The fuzzification is performed with real parameter \( n > 1 \).

\[
\mu_{jk} = \frac{1}{\sum_{i=1}^{C} \left( \frac{d_{jk}}{d_{ik}} \right)^{n-1}}
\]  

(3)

The FCM clustering algorithm is generally an iterative process, it is described as follows,

1. Choose \( n (n > 1) \) : The membership function values \( \mu_{jk} \) are initialized, \( j = 1, 2, ..., m; k = 1, 2, ..., c \).
2. Compute the cluster centers \( x_k \), \( k = 1, 2, ..., c \), according to Eq. (5).
3. Compute the Euclidean distance \( d_j \), \( j = 1, 2, ..., m; k = 1, 2, ..., c \).
4. In an each iteration, the membership function is updated \( \mu_{jk}, j = 1, 2, ..., m; k = 1, 2, ..., c \), based on the Eq. (6).
5. If it is not converged, go to step 2.

B. Hybrid encryption algorithm for improving the security

The data needs to transfer through the network is given as the input to the HEA for encrypting the data. Dual RSA and MD5 have been introduced in this HEA-AS method for improving the confidentiality of the network. The public and private key exponents are shared by using dual RSA. It leads to decrease the memory requirements to store both keys, because both the key exponents are similar to each other. The dual RSA is also utilized to decrease storage requirements. Here, the combination of the dual RSA and MD5 is developed for the best performance of the hashing function security. These encrypted features are transmitted from source node to the destination node (i.e., BS) system. The key generation and encryption process of the HEA-AS method is depicted as follows:

a. Process of Dual RSA key generation

Input: Create or select large random prime numbers.
Output: Public Key \( (e,n_1,n_2) \) and private key \( (d_1,d_2,p_1,q_1,p_2,q_2) \).

1. Select 4 different prime numbers such as \( p_1, p_2, q_1 \) and \( q_2 \)
2. Calculate the modulus \( n_1 = p_1 \times q_1 \) and \( n_2 = p_2 \times q_2 \)
3. Calculate the \( \phi_1(n_1) = (p_1 - 1) \times (q_1 - 1) \) and \( \phi_2(n_2) = (p_2 - 1) \times (q_2 - 1) \)
4. Select the public exponent as an integer \( e \) such that \( d_1 < e < \phi(n_2) \) and \( \gcd(e, \phi(n_1), \phi(n_2)) = 1 \).
5. Calculate the private exponent \( d = e^{-1} \mod \phi(n) \).
6. Public key \( = (e, n_1, n_2) \)
7. Private key \( = (d_1, d_2, p_1, q_1, p_2, q_2) \).

b. Dual RSA with MD5 encryption algorithm

Input: Plain text and public key \( (e, n_1, n_2) \) from dual RSA
Output: Cipher text.

1. Receive the authentic public key \( (n, e) \) from the key generation algorithm.
2. Compute \( X = F^e \mod n \) (i.e., \( n = \gcd(n_1, n_2) \))
   \[ Where, \ F^e \ is plain text. \]
3. \( Y = MD5(X) \)
4. Deliver the cipher text \( (Y) \) to the source node of the network.

This Dual RSA with MD5 encryption algorithm gives the encrypted values of plain text to the source node.

C. Prevention of malicious node in the ACO based routing

The data encrypted by using the HEA algorithm is transferred through the optimal path identified by using the ACO. The detection of an optimal path using ACO considers the node id, distance and residual energy of the cluster head. Moreover, the optimal path from the ACO is verified by AS scheme whether it has a malicious node or not. If the obtained path has malicious nodes, the ACO routing is again initiated to identify the optimal path without malicious nodes.

1. The CHs present in the network transmits their route information to the nearer CHs. The route information of CH contains node id, distance and residual energy. Then these CHs are saved the information in the routing table.
2. In the ACO based routing, the ant has located in each CH in the regular time intervals for identifying the route from the source to the destination. The ant placed on the CH identifies the next hop node by considering the probability function which is expressed in Eq. (4).

\[
P_o = \frac{\left( \tau_0 \right)^\alpha \left( \eta_0 \right)^\beta}{\sum_{j \in N} \left( \tau_j \right)^\alpha \left( \eta_j \right)^\beta}
\]  

(4)

Where the information of pheromone from source CH node to the next CH is denoted as \( \tau \) and the heuristic
Secured Data Transmission and Malicious Node Detection in Wireless Sensor Network

This pheromone information $(\tau_{ij})$ is represented in the Eq. (5).

$$\tau_{ij} = \frac{1}{d_{ij}}$$

(5)

Where, $d_{ij}$ is the Euclidian distance between source sensor $i$ and its associated CH.

The distance from one node to another node is computed as Euclidian distance which is expressed in Eq. (6).

$$d_{ij} = \sqrt{(S(i).xd - s(j).xd)^2 - (S(i).yd - s(j).yd)^2}$$

(6)

The heuristic information specifies the node’s energy level that is expressed in Eq. (7).

$$\eta_{ij} = \frac{E_{residual} - E_{residual}}{E_{residual} - E_{initial}}$$

(7)

Where, the initial and remaining energy of the node is denoted as $E_{residual}$ and $E_{residual}$ respectively. The relative weight of the heuristic value and pheromone trail are regulated by using two parameters such as $\alpha$ and $\beta$.

1. The sensor node with high amount of probability is selected as next hop node in the threshold based data transmission. The next hop node is selected in the path from the CH to BS.

b. Malicious node detection using acknowledgement scheme

The acknowledgement scheme of the HEA-AS method is divided into two phases, namely detection phase and the broadcast phase. The malicious nodes are identified in the detection phase and the information about the malicious node detection are transferred throughout the network. The detection phase of the HEA-AS method is divided into three different modes such as acknowledgement mode, conviction mode and confirmation mode. The acknowledgement phase provides the acknowledgement like whether the packets send and received without loss or not. The packet loss may happen when the transmission path identified from the ACO has malicious nodes. In that case, the conviction mode is used for discovering the node which drops the packets in the respective path. The outcome of the conviction mode is validated by using the confirmation mode. The process of two phases of acknowledgement scheme is given as follows:

1. If there is no acknowledgement is sent by the BS, then the conviction mode is initiated.
2. In the same route, a one hop request packet is transferred to the next hop sensor.
3. The next hop sensor will give acknowledgment signal when the node is a normal sensor. Then this process is continued for the next pair of nodes in the transmission path.
4. The misbehavior report is created when the next hop node doesn’t send any acknowledgment and then confirmation mode is initiated.
5. The confirmation mode validates the outcome of the conviction mode. Then the source node initiates the broadcast phase.

b. Detection Phase

1. At first, the sensors in the optimal path from the ACO is fixed into acknowledgement phase.
2. Then the source node sends the data packet and it waits for the acknowledgment signal from the BS.
3. If it receives the acknowledgement, then there are no malicious nodes presented in the optimal path. The data packets can transfer through the optimal path from the ACO.

Table I. Simulation parameters

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area</td>
<td>$100m \times 100m$</td>
</tr>
<tr>
<td>Nodes number</td>
<td>100</td>
</tr>
<tr>
<td>Antenna model</td>
<td>Omni-directional</td>
</tr>
<tr>
<td>Clustering algorithm</td>
<td>FCM</td>
</tr>
<tr>
<td>Routing algorithm</td>
<td>ACO</td>
</tr>
<tr>
<td>Encryption algorithm</td>
<td>Dual RSA-MD5</td>
</tr>
</tbody>
</table>

IV. RESULTS AND DISCUSSION

The simulation of the HEA-AS method against a scenario with the attacks modelled is performed using the network simulator 2.35. The nodes in the network are verified by using acknowledgement scheme like whether it is malicious node or not. Then the transmission over the WSN is secured by using the combination of dual RSA and MD5 algorithm. There are 100 nodes considered for the simulation and these nodes are deployed in the area of $100m \times 100m$. The channel characteristics and node’s parameters are specified in the Table 1. The IEEE 802.11 (i.e., MAC protocol) is considered in this HEA-AS method and this IEEE 802.11 utilizes the traffic sources of continuous bit rate. The performance measure is calculated in this HEA-AS method are packet delivery ratio, packet loss rate, end to end delay, throughput and network lifetime. The performance of the HEA-AS method is compared with existing method CTCM [16] for evaluating the effectiveness of the HEA-AS method.
A. Performance analysis

The performance comparison of the HEA-AS method with existing method CTCM [16] is given as follows:

a. End to end delay

It is defined as the amount of time taken for transmitting the data packets from the source node to the destination.

Fig. 2 shows the comparative analysis of EED for the HEA-AS method and CTCM method [16]. The graph of Fig. 2 is plotted as number of nodes verses time taken for the data transmission. From Fig. 2 its concluded that the EED of HEA-AS method is less when compared to the CTCM [16]. Because the ACO routing used...
in the prosed method gives an optimal path by considering two different factors CH’s residual energy and distance.

b. Throughput

The throughput is defined as the amount of successful packets transmitted over the optimal path generated by the HEA-AS method. The throughput of the network is calculated as bits per second (bit/s or bps).

Fig. 3 shows the comparative analysis of throughput for the HEA-AS method and CTCM method [16]. The graph of Fig. 3 is plotted as number of nodes verses amount of packets transferred to the BS. The throughput of the HEA-AS method is high when compared to the CTCM [16]. The reason behind the higher throughput is that the HEA-AS method uses the AS based malicious node detection in ACO routing and also the residual energy of the nodes is considered in the ACO routing. The packets drop in the network is avoided by rejecting the dead nodes in the data transmission.

c. Network lifetime

The lifetime of the network is defined as the time at which the first node dies in the network. The dead node is identified by monitoring the remaining residual energy of each node at every time.

Fig. 4 shows the comparative analysis of network lifetime for the HEA-AS method and CTCM method [16]. Fig. 4 is plotted as number of nodes verses time at which first node die through the network. From the Fig. 4 shows that the lifetime of the HEA-AS method is higher than the CTCM. The residual energy of the nodes in the HEA-AS method is preserved by identifying the optimal path for data transmission.

d. Packet deliver ratio

PDR shows the data transmission reliability. PDR is the ratio of actual packets successfully received by the sink to the total packets sent by the source in a network.

Fig. 4. Comparison of network lifetime
Fig. 5 shows the comparative analysis of PDR for the HEA-AS method and CTCM method [16]. The graph of Fig. 5 is plotted as number of nodes verses percentage of successful transmission of packets. The PDR of the HEA-AS method is high when compared to the CTCM method [16]. The PDR rate is maximized by monitoring the residual energy at every time to avoid the dead nodes in the data transmission path.

e. Packet loss rate

The PLR is defined as the ratio of the number of lost packets to the total number of sent packets.

Fig. 6 shows the comparative analysis of PLR for the HEA-AS method and CTCM method [16]. The graph of Fig. 6 is plotted as number of nodes versus percentage of dropped packets through the network. From the Fig. 6 conclude that the PLR of HEA-AS method is slightly decreased than the CTCM method [16]. The PLR is minimized by avoiding the malicious nodes and dead nodes in the data transmission path.

V. CONCLUSION

In this paper, the FCM based clustering is used to separate the network into a number of clusters. Then the respective CH is selected from the cluster
Secured Data Transmission and Malicious Node Detection in Wireless Sensor Network

to gather the data from its cluster members. The communication among the network is secured by combining the Dual RSA with MD5. This combination of encryption algorithms leads to increase the integrity and confidentiality of the network. Moreover, the ACO routing is used for transmitting the information from source node to BS. Here, the AS is integrated in the ACO routing into detect and prevent the network from attacks due to the malicious nodes. The performance of the HEA-AS method is compared with CTCM method. From the comparison, it is evident that the proposed HEA-AS method provides higher performance when compared to the CTCM method. Furthermore, the network lifetime of the WSN can be improved by using the optimum malicious node detection scheme and effective optimization algorithms for route generation.

REFERENCES


AUTHORS PROFILE

Deepak N. Biradar is pursing Ph.D under VTU, Belagavi, from 2015. Currently serving as Assistant professor in the Department of CSE at Lingaraj Appa Engineering College, Bidar from 2011 to till Date. His area of interest is Wireless sensor Networks, Fuzzy Logic.

Dr. T.S. Vishanath is presently working as Professor in the dept. of ECE, BKTIT Bhalki, Karnataka. Affiliated under VTU Belagavi. The author has published many papers in the area of signal processing, image processing and Sensor Networks. The author is a fellow member of ISTE, IETE, IEL. And life member of IJEEE.