Journal of Trust Management

RESEARCH Open Access

LDAT: LFTM based data aggregation and transmission protocol for wireless sensor networks

Mukesh Kumar* and Kamlesh Dutta

* Correspondence:
mukeshk.chawla@gmail.com
Department of Computer Science & Engineering, National Institute of Technology, Hamirpur, Himachal Pradesh 177005, India

Abstract

Wireless sensor network consists of a large number of resource constrained sensor nodes. These sensor nodes communicate over wireless medium to perform a variety of information processing functionality. Due to broadcast nature of wireless medium, security is one of the major concerns and overlapping sensing range of sensor nodes results in redundancy in sensing data. Moreover, a large amount of energy is consumed by the base station to process these redundant data. To conserve energy and enhance the lifetime of sensor nodes, redundancy is eliminated at intermediate nodes by performing data aggregation. Wireless sensor networks are generally deployed in untrusted and hostile environments which results in compromised nodes. Thus, security and reliability of the transmitted data get reduced. Compromised nodes can inject false data, drop all the data, selectively forward data to an attacker, copy legal nodes to join routing paths, and disrupt data transmission during the data aggregation operation. In this paper, a novel scheme for data aggregation based on trust and reputation model is presented to ensure security and reliability of aggregated data. It will help to select secure paths from sensor nodes to the base station; thereby the accuracy of aggregated data will be increased significantly. Simulations show that the proposed protocol LDAT has less energy consumption and more accuracy as compared to some existing protocols which are based on functional reputation.

Introduction

Data Aggregation is one of the methods to reduce the communication burden in which a sensor node naming data aggregator processes and aggregates incoming data before passing it to its neighbour node. Data Aggregation is the essential technique to achieve energy efficiency by reducing data redundancy and optimize the bandwidth usage. Obviously, with energy consumption the security of WSN must also be taken into consideration, when they are deployed in an insecure environment. Several security mechanisms can be used to keep the Data Aggregation process secure such as cryptography, key management, MAC (Message Authentication Code) mechanism.

Pure cryptography cannot provide proper security for Data Aggregation because cryptography cannot provide defense against node capture attack which results in compromised nodes [1]. In order to achieve more security in Data Aggregation, reputation and trust system are used for monitoring network activities and events. Trust and reputation models are utilized to detect, collect, process the sensor's recent behaviors and then calculate the node's trustworthiness for specific application. Various activities of the sensor node are concerned for evaluation of trustworthiness of that node such as data collection, data transmission, aggregator selection and routing path selection [1].

Trust models define the method with which the trust information and trustworthiness of each node are obtained. The aim of using trust models with data aggregation is to improve security and increase the lifetime of the network. The main goal of using the Trust and Reputation system is to defend the network against compromised nodes and remove these nodes from further participation in Data Aggregation process [2].

Problem of managing trust and reputation model during Data Aggregation in Wireless Sensor Networks (WSN) in an efficient, accurate and robust way has not been completely solved yet. Till now various Trust and Reputation models have been introduced and they all are varying from mathematical approach to biological approach. This paper first discusses performance analysis of various trust and reputation models for WSN. Simulation result shows that LFTM (Linguistic Fuzzy Trust Model) trust model provides more accuracy even if the malicious nodes are large in number and it consumes less energy.

The objective of this paper is to provide the comparison of various trust mechanism with short summarization of trust methodologies in WSN which can provide a high level of security taking into account accuracy, average path length leading to trustworthy sensors and energy conservation. LDAT (Linguistic Fuzzy Trust based Data Aggregation and Transmission) is proposed in this paper that evaluates the trustworthiness of sensor nodes during Data Aggregation to improve the reliability and accuracy of aggregated data. In protocol LDAT, security of the Data Aggregation process is ensured by selecting the trusted data aggregator using Linguistic Fuzzy Trust mechanism.

Trust is one of the security mechanisms which tries to detect inside attacks and constructs a self healing WSN. Trust and Reputation System helps to maintain a minimum security level between the entities of distributed systems for interactions or transactions. These entities in a WSN are data aggregator node, normal node and base station. The general purpose of using Trust System is to enhance the security of network, but there are other applications of Trust Systems such as access control [3], Data Aggregation [4], routing [5] and intrusion detection in Wireless Sensor Network. A Trust and Reputation model is generally composed of five components [6, 7] such as gathering information, scoring and ranking, selecting entities, having transactions and reward or punishment. In this paper LFTM trust model is modified to aggregate the data and to provide security and reliability to the aggregated data.

Trust terminologies

i) Node

   The Node is a basic individual unit in sensor network [7].

ii) Cluster

The Cluster is a group of nodes in a network. Clustering is preferred as it reduces the communication overhead over the network [7].

iii) Trust

The trust in general is interpreted as a belief or some subjective probability assigned to the node in the network. Trust is a subjective opinion in the reliability of other entities, including reliability of data, connectivity of path, processing capability of node and availability of service etc. [8].

iv) Reputation

Concept of reputation is considered as a close relevance measure to evaluate trust based on the recommendation from other participants in the network but according to authors [8], these two terms are clearly different as illustrated by following statements.

- I Trust you because of your good reputation.
- I Trust you despite your bad reputation.

v) Trust values

Trust values provide various methods of evaluation. Generally, there are two types of trust values: continuous and discrete. In continuous type values, there is varying range of trust example [−1, 1]. The discrete trust value may be depicted by an integer number or discrete value with labels rather than numbers. Some algorithms use values ranging from negative to positive [9].

Trust and reputation models in WSNs

In this section a brief overview of various trust and reputation models for WSN is given. These models are then compared under simulation environment. The model which is best amongst all the models will be modified as per our requirements and will be used in the proposed protocol.

Peer

Peer Trust model [8, 10] is a dynamic peer-to-peer trust and reputation model, initially aims to evaluate the trustworthiness, or goodness of participating peer and to combat the selfish, dishonest and malicious peer behaviour. The Peer Trust System computes the trustworthiness of a peer by the average feedback given by the scores of the feedback originators. Peer calculates trust score over five factors in a distributed manner, namely: 1) the feedback a peer retrieves from others; 2) the feedback scope, or field (number of transactions); 3) the credibility factors of the source; 4) the transaction context factor addressing the criticalness of transactions; as well as 5) the community context factors interpreting related characteristic. The limitation of this approach is that the computation of convergence rate in large scale P2P (Peer to Peer) system is not provided [9].

Power

Power Trust [11] is a robust and scalable P2P reputation system to control power-law feedback characteristics. Power Trust model dynamically selects a small number of power nodes that are most reputable using a distributed ranking mechanism. Power Trust significantly improves global reputation accuracy and aggregation speed. The major building blocks of Power Trust are a trust overlay network (TON) built on top of all peers in P2P system; regular random walk module that supports the initial reputation aggregation and look-ahead random walk (LRW) module that updates the reputation score. This module also works with a distributed ranking module to identify the power nodes. Power Trust system [11] is robust to resist malicious peers and high scalability to support large-scale P2P applications.

BTRM (Bio-inspired trust and reputation model)

BTRM-WSN [6] is a bio-inspired trust and reputation model for WSNs aimed to achieve most trustworthy path leading to the most reputed node in a WSN. It is based on the bio-inspired algorithm of an ant colony system, but due to the specific restrictions and limitations found in WSNs, the ACS (Ant Colony System) cannot be directly applied. Some adaptations, therefore, have to be made. In the improved BTRM [12], each sensor node in the network contains pheromone traces. Pheromones are the hormones which secrete on the move and determine the probability for an ant to select a path as well as the sense the path leading to a solution. This algorithm has mainly three steps namely, gathering information, having transaction and last step is giving rewards or punishment.

LFTM

Linguistic Fuzzy Trust model (LFTM) [13] deals with linguistic fuzzy labels, which are closer to the human way of thinking and also uses the fuzzy reasoning. This model keeps the accuracy of the underlying bio-inspired trust model and the level of client satisfaction while enhancing the interpretability of the model and thus making it closer to the end user. The Linguistic fuzzy logic and fuzzy reasoning provide the framework for knowledge representation, model transparency and inference for a trust model for distributed network system. An ant-colony optimization will be guided using such LFTM. This model is able to provide a platform that achieves very high levels of client satisfaction.

Related work

This section discusses the related work to introduce the trust based system for securing Data Aggregation in WSN. Discussion is started with the technique [14] in which authors proposed a framework for secure information aggregation (SIA) in large network. In this framework random sampling mechanisms and interactive proofs help the user to verify that the answer given by aggregator is a good approximation of the true value even when the aggregator and a fraction of the sensor nodes are corrupted. A separate secret key is shared with the home server and an aggregator that enables message authentication and encryption of message. The protocol [14] also proposed a forward secure authentication approach that ensures that if an attacker corrupts a sensor node at any point of time, it will not be able to modify the previous data which the sensor node has recorded locally. Ganeriwal and Srivastava [15] proposed reputation based framework in which nodes maintain the record of reputation of other nodes and use this information to evaluate their trustworthiness. This provides a generalized approach for detecting the malicious or faulty nodes in the network. In this technique, for representing and updating of reputation, Bayesian formulation specifically beta reputation system is used. In the next technique [16], authors proposed a trust based framework in which Kullback-Leibler (KL) is used to evaluate the trustworthiness of sensor nodes in which compromised nodes are detected with the help of unsupervised learning.

SAT [17] i.e., Secure Aggregation Tree is a method that detects and prevents cheating in the network. This method does not use any cryptographic operation; however, detection of cheating is based on the topological constraints in the aggregation tree. This method is different from other proposed methods as it is based on cheating detection instead of persistent data authentication. With topological constraints in SAT, each node can hear all messages sent to its father node and message sent by a father node to its grandfather node, that help to check whether father node performs the Data Aggregation correctly or not. If the father node sends aggregated value which is different from the correct aggregation value, then the node will raise an alert. On receiving the alert message by neighbouring nodes they all check whether the cheating node is its father node or not. If yes, then weighted confidence is evaluated. If the weighted confidence value is larger than a predefined threshold value, then the father node is assigned to be cheating node and a detection confirmation message is broadcasted within a given hop limit. After that, the node receiving the detection confirmation message will use recovery mechanism to avoid using the compromised node. During Data Aggregation, a metric to represent the degree of belief is generated for defining the uncertainty in the aggregated result. This framework effectively measures the uncertainty in both data and aggregation result.

In centralized trust system, there is centralized infrastructure of trust that keeps the reputation values at centralized authority and controls the system in the network. While concerning the characteristics of WSN, centralized trust based system are not feasible because with the increase in the number of sensor nodes in the network, the scalability and expandability of the network is not supported by centralized trusted centre. Due to this factor, decentralized trust based systems are being developed and used in wireless sensor networks. Trust development system in the technique [18] uses combined trust values (CTV) to favour packet forwarding for each node without using any centralized infrastructure. In this protocol each node has CTV value that evaluated on the basis of three factors as identification, sensing data and consistency.

Recently trust based systems that are used in WSNs are classified into five categories based on their different applications that are generic, routing, access, location and aggregation [19]. Recently Trust development systems RDAT [4], iRETDA [1] are proposed for wireless sensor networks that use the concept of functional reputation. In general reputation system, reputation is computed over all actions of sensor nodes. Functional reputation system detects the compromised nodes that cover its bad actions with respect to one function by behaving well for other functions. This protocol [4] is based on the concept of functional reputation. The evaluation of action of sensor nodes by data aggregators based on the respective functional reputation of nodes increases the reliability and accuracy of trust system. The main focus of this protocol is to mitigate the effect of compromised nodes on Data Aggregation. The functional reputation of the sensor node is represented by the beta distribution [4]. Protocol evaluates the functional reputation values for each sensor node by evaluating respective functions separately. Three functional reputations are concerned to evaluate the trustworthiness of sensor nodes that are sensing, routing and aggregation. The Tiny OS 2.0 (TOSSIM) simulator was used for performance evaluation of RDAT Protocol. It shows that using the functional reputation concept of reputation system is more effective than using the overall reputation concept for evaluating the trustworthiness of sensor nodes. The next protocol iRTEDA [1] provides information about the residual energy and link availability to assess the trustworthiness and reliability of sensor nodes based on the observation of neighbouring nodes. In this paper, the comparison of the proposed protocol is done with both RDAT [4] and iRTEDA [1].

Preliminaries

System model

In the proposed work the hierarchical cluster architecture is considered to construct WSN with densely deployed sensor nodes. It is assumed that each cluster is performing their operation independently and very few overlapping areas would be sensed between the clusters. Due to a large number of sensor nodes in each cluster, there is overlapping of sensing ranges and event by multiple sensor nodes. Hence, to remove redundant data that are aggregated of correlated data at the neighbouring sensor node is required. Each cluster has a cluster head that acts as aggregator, to aggregate data from their neighbouring sensor nodes. In the WSN, an attacker can compromise sensor node and after that attacker will get all the information about the node. Every sensor node in the network is capable of doing Data Aggregation. The role of data aggregator is rotated among the sensor nodes of the network to balance the energy consumption of nodes. Figure 1 shows the cluster structure of the proposed work.

Data aggregation techniques, explore how the data are to be routed in the network as well as the processing method that are applied to the packets received by a node. The designated nodes are called data aggregator and the process is data aggregation, which is shown with a model through which we understand the difference between communication with data aggregation and non data aggregation as shown in Fig. 1.

In communication with non data aggregation, every node sends and receives messages on the single transmission path. They consume more energy, bandwidth and time. These three parameters are important in wireless sensor network and their values should be low when we transfer any message between the nodes. In the data aggregation process each low level node sends their data to its above level node, that above level node become an aggregator that aggregates all the data coming from its lower level node and then send to its above level node and so on as shown in Fig. 1. In this mechanism the three parameters (energy, bandwidth, time) considered in communication without data aggregation are low.

In a particular wireless sensor network whenever there are data readings to send by sensor node, it passes to its upper level nodes designated as aggregator nodes. It can be seen from the Fig. 2, when there is no data aggregation and every sensor node has to send the data to base station, the number of message transfers is 22. In case, aggregation process is applied it is significantly reduced to 9 messages

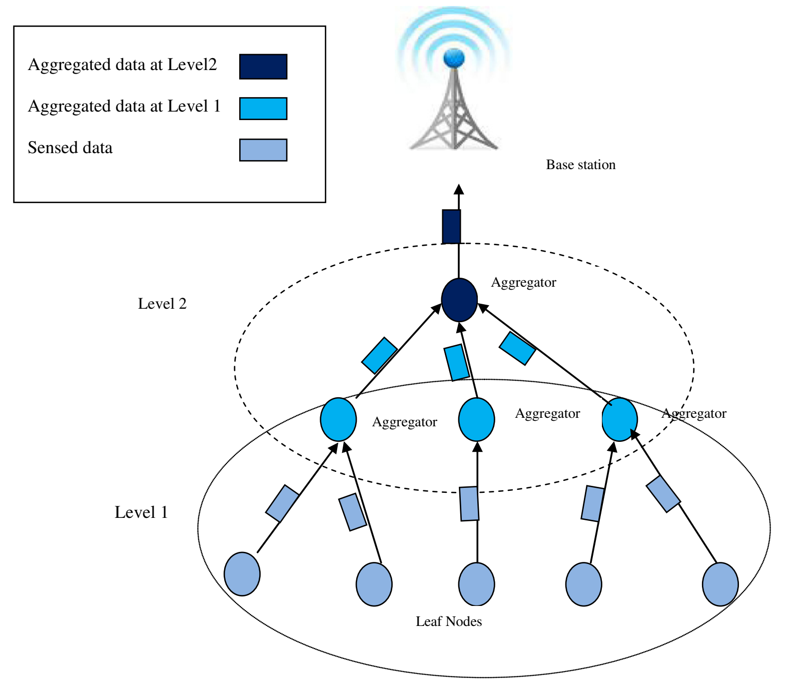


Fig. 1 Proposed Architecture for a WSN

which is depicted from the Fig. 3. Thus, the communication message transfers are roughly reduced to 41 % as shown in Table 1. It can be observed that how important is data aggregation in case of wireless sensor network and why it should be protected from various attacks.

mLFTM (Modified LFTM Trust Model)

The main objective of the proposed work is to assess the application of linguistic fuzzy sets and fuzzy logic to several concepts and to enhance the existing trust and reputation model [13]. This model will use the inference power of fuzzy logic as imprecise

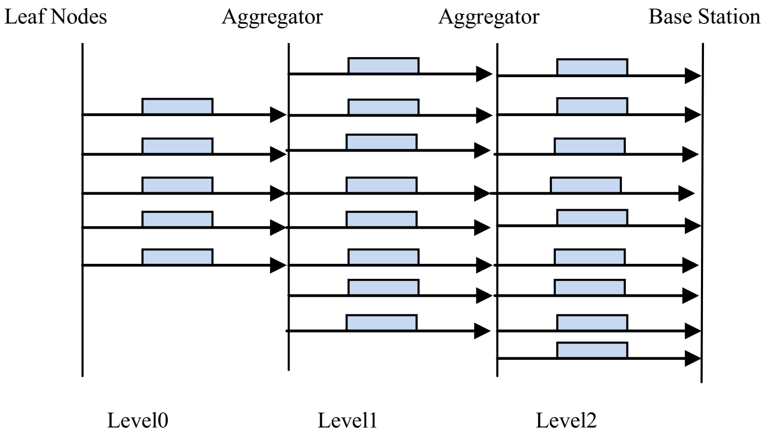


Fig. 2 Communication without data aggregation

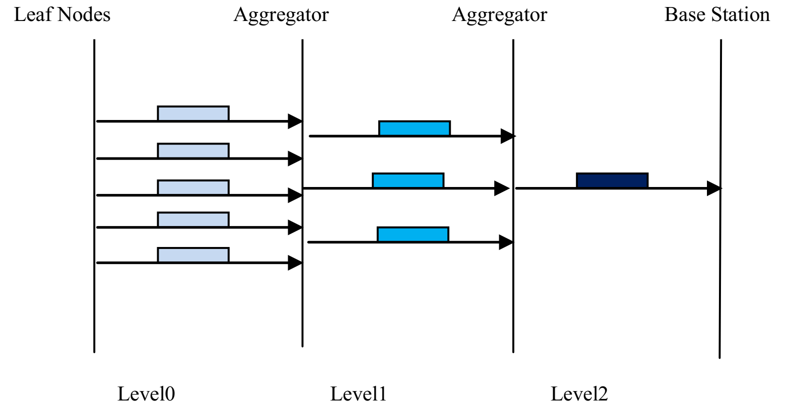


Fig. 3 Communication with data aggregation

dependencies between original requested aggregated data and actually received one and also use the representation power of linguistic labeled fuzzy sets e.g., satisfaction of base station or goodness of the aggregator.

Assumptions: The modified trust model is based on some previous assumptions as mentioned in LFTM and BTRM-WSN and some new assumptions as discussed below.

Node Terminology: Original LFTM and BTRM models assumed that in the network there are two subsets of sensor nodes: client nodes which request some services and server nodes which provide some services. In modified model every node is capable of providing the services to all other nodes. In modified model there are five types of sensor nodes. Their roles are revised in new model.

a) Normal Sensor Node. In case of LFTM [13] known as client node. These nodes have data to be sent to base station through aggregator nodes. Algorithm is executed on these nodes in order to find good server/aggregator.
b) Aggregator Node: In case of LFTM model, these nodes were called server nodes. These are the nodes which aggregate the data coming from different sensor nodes.
c) Malicious Node. These are the nodes which do malicious activity. In LFTM model malicious activity was not defined. But in this proposed model we have defined malicious activity as packet loss (0–40 %).
d) Intermediate Node. In case of LFTM model these were known as relay nodes. These nodes just forward the data as received.
e) Idle Sensor node: These are the nodes which remain in sleep mode until asked to send the data.

Table 1 Difference between Data Aggregation and Non Data Aggregation

Non Data Aggregation		Data Aggregation	
Leaf Nodes	5 messages	Leaf Nodes	5 messages
Level 1	8 messages	Level 1	3 messages
Level 2	9 messages	Level 2	1 message
Total	22 messages	Total	9 messages

Topology: In our proposed work we have considered dynamic topology.
Node behaviour: Every node is capable of knowing only about its neighbours.
Steps for mLFTM which in turn based on BTRM-WSN

1. *Gathering information*
   When the algorithm is launched, a set of artificial ants are deployed over wireless sensor network. These ants leave some pheromone traces throughout the paths they travel. Main goal is to find the most trustworthy aggregator node required by the sensor node executing BTRMWSN. To do so, they follow the pheromone traces left by previous ants. Thus, the greater the pheromone trace a specific path has, the more suitable such route is to be selected as the one leading to the most reputable node.
2. *Scoring and ranking*
   Once the ants have found a path including a trustworthy aggregator to the base station, a score has to be given to each of those paths. Such assessment is performed through the following expression shown in Eq. 1.

$$Q(S_k) = \left(7_k/length(S_k)^{PLF}\right)\%A_k \tag{1}$$

   Here S_k is the path returned by ant k,7_k is average pheromone traces of such path, *PLF* is a path length factor, $\%A_k$ is the percentage of ants that give the same path as ant k. After that the path S_i with the highest value of Q (S_i) is selected by BTRM to have transaction with LFTM.
3. *Aggregator Selection*
   The path S_i with the highest value of Q (S_i) is selected by BTRM-WSN as the one leading to the base station through the most trustworthy aggregators in the network.
4. *Aggregation and Satisfaction Level Computation*
   The sensor node sends the data to the selected aggregator node that will forward the data to another aggregator nodes or base station depending on its goodness. Base station, then evaluates the received results and computes its satisfaction with the performed aggregation.
5. *Rewarding and Punishing*
   If the base station was satisfied with the received aggregated results, reinforcement in terms of pheromone additions to the path leading to the final service provider is carried out. Otherwise, if the aggregator has cheated, a punishment in terms of pheromone evaporation is carried out. One modification we have added to reward and punishment criteria. We do not either promote such path or downgrade the whole path. Instead, the node which has cheated will be given a reward or punishment. It is unfair to give punishment to whole path if just one node is found to misbehave.

Figure 4 depicts the flow of our approach, emphasizing those steps where we actually applied linguistic fuzzy sets and fuzzy logic. Such steps are as follows:

1. The trust and reputation model *m*BTRM-WSN selects the aggregator node to have a transaction with.
2. Such aggregator node has a perceived certain goodness ('very high', 'high', 'medium', etc.).

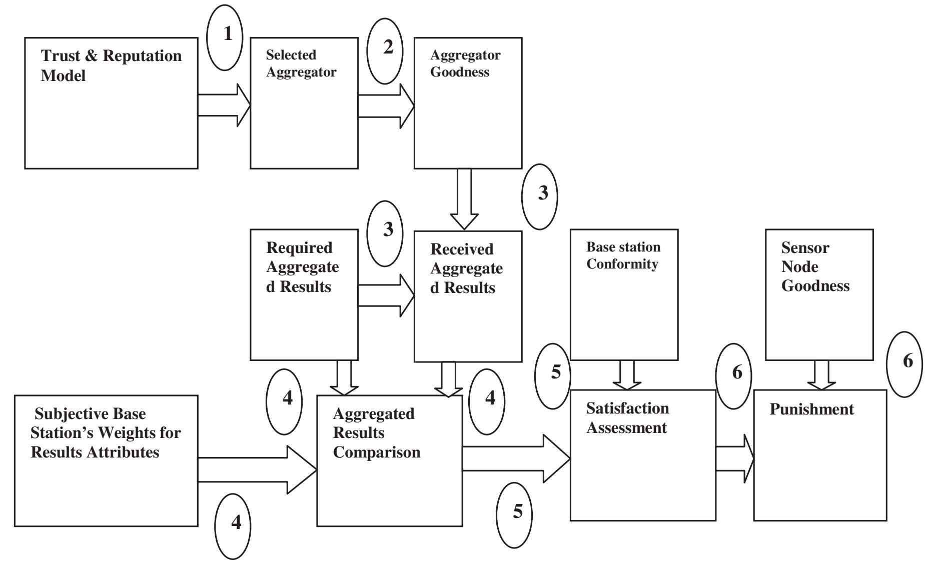


Fig. 4 Modified Linguistic Fuzzy Trust Model steps

3. According to the required aggregated result attributes and the aggregator goodness, the aggregator node provides a better, worse, or equal aggregated results than the expected results.
4. Both the required aggregated results and the actually received one are compared using certain subjective and sensor node dependent weights of the aggregated results attributes.
5. The satisfaction is assessed by means of the aggregated result comparison performed in the previous step and the base station conformity.
6. Finally, the punishment level is determined by the satisfaction with the received aggregated results, together with its goodness.

LFTM uses a strong fuzzy partition with the property that any measured value of the variable would have some membership to at least one linguistic fuzzy set and at most of two, so full coverage is obtained. With the application of linguistic fuzzy sets and fuzzy logic in LFTM each server has a certain goodness, i.e., 'very high','-high','medium' etc. Service in the Trust model is composed with four attributes: price, cost, quality and delivery time. According to required service attributes and server goodness, the server provides the service. After receiving the service from the server, the client compares its attribute individually with the corresponding attributes of the requested service.

A client node in the model gives some subjective weight to each service property and the weighted aggregation of two service properties is performed to obtain final result. After that client satisfaction is evaluated by comparing the client conformity with received service. Finally, the client satisfaction together with the client goodness, will determine the reward/punishment level of the selected server.

In the current work, a strong fuzzy partition will be used to define the underlying fuzzy sets (that will be linguistically labeled). A strong fuzzy partition has the following properties being S_i the fuzzy sets defined over the domain D and x a value of such domain:

$$\begin{aligned} &\forall i, \exists x \in D, \mu S_i(x) = 1 \\ &\forall x \in D, \exists i, j \, \forall k \;\; i \neq j, k \neq i, k \neq j, \end{aligned} \tag{2}$$

$$\begin{aligned} \mu Si\,(x) + \mu S_j\,(x) &= 1 \\ \mu S_k(x) &= 1 \end{aligned} \tag{3}$$

The first expression in Eq. 2 ensures normality. The second expression in Eq. 3 states that any particular value of the domain can belong, at most, to two different fuzzy sets (S_i or S_j) and that the addition of the membership values for any given value of the domain is equal to one linguistic label L_i will be associated with each defined fuzzy set S_i.

Advantages of mLFTM over existing LFTM model

The main advantage of *m*LFTM is less energy consumption due to the process of data aggregation. In mLFTM model dynamic topology is used as in earlier LFTM model. In our proposed algorithm, we have modified reward and punishment criteria. In existing LFTM protocol, punishment was given to entire path, but in a proposed algorithm only the malicious node will be punished.

LDAT: LFTM based data aggregation and transmission protocol

In this section a LFTM based Data Aggregation and Transmission protocol (LDAT) is proposed which is based on linguistic fuzzy trust mechanism for distributed networks. The proposed protocol is divided into two sub protocols TDA (Trust based Data Aggregation) and RDT (Reliable Data Transmission).

Reliable data aggregation

In protocol TDA, data aggregation is periodically performed in definite time intervals. In each data aggregation session, reliable data aggregation is achieved in two phases. In the first phase, before transmitting data to data aggregators, each sensor node A_j using *m*LFTM trust model. If trustworthiness of A_j is below to predetermined threshold, then N_i will not send data to A_j. To achieve this, N_i sends this data to the base station along with a report indicating A_j may be compromised. Based on the number of reports about A_j over the time, the base station will decide that A_j is a compromised node and it should be removed from the network. In the second phase of data aggregation session, Trust based Data Aggregation (TDA) algorithm is run by data aggregators. Algorithm TDA depends on Q (S_i) value to mitigate the effect of compromised sensor nodes on aggregated data.

Table 2 shows the packet structure used in our protocol. Each node in the network has Node ID that is unique and each data aggregator node has table that maintain the list of nodes that are in its cluster. DEST and SRC field in the packet, determine the source address and destination address respectively. The byte size of each field is indicated in the enclosed parentheses.

Table 2 Packet structure of protocol LDAT

Node ID (3)	Seq No.(4)	DEST(3)	SRC(2)	LEN(2)	DATA(0..56)

Each sensor node maintains the table containing the reputation value of its aggregator and this table is exchanged among other sensor node of the same cluster. During Data Aggregation the following algorithm is run by data aggregator node.

Algorithm: TDA (TRUST based DATA AGGREGATION)

Input: *Data aggregator A_j, A_j's neighbouring nodes (N_1, N_2...N_n), Q values of neighbouring node computed by $A_{j.}$*

Output: *Aggregated data D_{agg}.*

Step 1: *A_j requests each N_i to send its data for Data Aggregation.*

Step 2: *$A_{j\ updates}$ Q (S_i) values of the nodes $N_{i.}$*

Step 3: *Sensor nodes (N_1, N_2...N_n). Transmit data (D_1, D_2, .D_n) to A_j.*

Step 4: *Aj weights data Di of sensor node.*

Step 5: *A_j aggregates the weighted data to obtain D_{agg}.*

The scheme is described as follows:

At Sender Node:

(1) Send request Message

(2) Create neighbour table using algorithm NT

(3) Select aggregator Node A on the basis of selection criteria (Residual energy)

(4) Verify trustworthiness of Aggregator A_j using proposed protocol.

At Aggregator Node:

(1) Send reply message M_{rep}.

(2) Apply aggregation function F (D_i) s.t.

F (D_1, D2, Dn) = D_{agg}. Where D_i is reading of the respective sensor node.

(3)Send D_{agg} to the next hop or base station, whichever is the nearest.

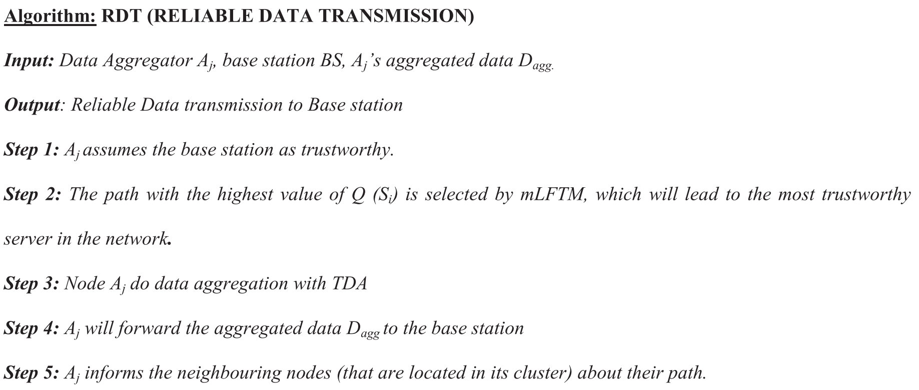
At Base Station:

(1) Trusted and aggregated data is received by the base station.

Reliable path selection

When a data aggregator node A_j needs to send the aggregated data D_{agg} to the base station, A_j selects the reliable path by executing mLFTM. The path with highest Q (S_i) value is selected for data transmission. Several different parameters have been proposed to select the reliable path in sensor networks [14]. However, these path selection methods do not consider the security aspect of the paths. In order to select the reliable and trustworthy path from data aggregator node A_j, periodically LFTM's ant colony algorithm returns the path with different trust values [20]. Algorithm RDT selects the path with the highest score. mLFTM trust and reputation model for WSNs aimed to achieve the most trustworthy path leading to the most reputable node in a WSN offering a certain service. It is based on the bio-inspired algorithm of ant colony system but, due to the specific restrictions and limitations found in WSNs, the ACS cannot be directly applied. Some adaptations, therefore, have to be made. In our model, for instance, every node maintains a pheromone trace for each of its neighbours. This pheromone traces will determine the probability of ants choosing a certain route or another, and can be seen as the amount of trust given by a node to other one. The heuristic values are defined as the inverse of the delay transmission time between two nodes (or the inverse of the distance between them). This algorithm [8] consists of the following steps:

1. Every ant adds the first sensor to its solution, which is always the client they are departing from.
2. Once every ant has left the client, this one waits until they come back.
3. The best solution found by all or some of the ants issued in the current iteration is compared with the global best solution and swapped if it is appropriate.
4. Finally, a pheromone global updating is performed over the links belonging to the global best path.



Algorithm: RDT (RELIABLE DATA TRANSMISSION)

Input: *Data Aggregator A_j, base station BS, A_j's aggregated data $D_{agg.}$*

Output*: Reliable Data transmission to Base station*

Step 1: *A_j assumes the base station as trustworthy.*

Step 2: *The path with the highest value of Q (S_i) is selected by mLFTM, which will lead to the most trustworthy server in the network.*

Step 3: *Node A_j do data aggregation with TDA*

Step 4: *A_j will forward the aggregated data D_{agg} to the base station*

Step 5: *A_j informs the neighbouring nodes (that are located in its cluster) about their path.*

In the above algorithm reliable data transmission will be performed whenever the data aggregator node A_j has the aggregated data D_{agg} for the base station. After getting the Q (S_i) value of path S_i, the path with highest Q (S_i) value is selected. After transferring the aggregated data to the base station data aggregator node will inform the neighbouring node about the path selected by it and its trust value.

Performance evaluation

In this section performance of various trust and reputation models is evaluated and compared. Trust model which is best among other models in terms of accuracy, energy comparison, and average path length is used in the proposed protocol. Then performance of the proposed protocol is compared with the existing protocols RDAT [4] and iRTEDA [1] with the help of TRMSim-WSN [21]. Parameters for evaluations are accuracy, average path length, energy consumption and packet delivery ratio. TRMSim-WSN allows the user to adjust several parameters to run the simulation as follows:

- *Percentage of client nodes*: The percentage of nodes that want to send message to other nodes in the network and ask for services in a WSN.
- *Percentage of relay nodes*

$$= (\text{No of relay nodes / Total no of nodes}) * 100. \tag{4}$$

- *Percentage of malicious nodes*:

$$(\text{No of adversaries nodes / Total no of nodes}) * 100. \tag{5}$$

- *Radio range*: A distance within which the nodes are able to sense each other. Other sensors within the range of a node can be considered as its neighbours.
- *Delay*

$$\mathrm{D_T} = \mathrm{N} / \mathrm{R} \tag{6}$$

  where $\mathrm{D_T}$ is the transmission delay, N is the number of bits, and R is the rate of transmission.
- *Number of executions*: The number of execution represents how many times the test runs.
- *Number of networks*: The number of WSNs simulated.
- *Min./Max. Number of sensors*: The minimum and maximum number of sensors in a random generated WSN.
- *Trust and reputation model*: Five trust and reputation models have been built into TRMSim-WSN 0.5: BTRM, Peer Trust Model, Eigen Trust Model, Power Trust, and LFTM.
- *Accuracy*: The selection percentage of trustworthy nodes :

$$(\text{Number of successful transmission/ total number of message transfers}) * 100. \tag{7}$$

- *Path Length*: The number of hops of the paths found by a trust and reputation system leading to the nearest trustworthy nodes.
- *Energy Model:* Consumption of the overall energy is sum of:
  1) Client nodes sending request messages;
  2) Server nodes sending response services;
  3) Energy consumed by malicious nodes, which provide bad services;
  4) Relay nodes which do not provide services; and
  5) The energy to run TRM executions.

This model is used to measure the energy of each sensor node. The energy consumed by each node is calculated as shown in Eq. 8.

$$E_T = E_R * K + E_T * K * L^2 \tag{8}$$

Table 3 Parameters for execution

NumExecution	100	% client	20 %				
NumNetworks	100	% Relay	5 %				
MinNumSensors	100	% malicious	{40 %, 60 %, 80 %, 95 %}				
MaxNumSensors	200						
Radio range	10	phi	0.01	No. of ants	0.35	rho	0.87
Niter	0.59	TraTh	0.66	PLF	0.71	alpha	1.0
q0	0.45	beta	1.0				

Where E_R is receiver's electronics energy and assumed equal 50, E_T is transmission energy of radio frequency (RF) signal generation and it is considered equal to 100, K is the number of bytes (packet size capacity of each node), L is the radio range of each node [19].

A topology where a WSN is composed of nodes with relatively high sensor activity is considered in this protocol. It is considered that some nodes are requesting generic services and some nodes are providing them. Various parameters considered for execution is shown in Table 3.

Evaluation of trust and reputation models

Evaluation of the TRM models is performed by TRMSim-WSN [21] which is a Java-based simulator aiming at providing an easy way to test a trust and reputation model over WSNs. It compares the model against other models over three main parameter accuracy, path length and energy consumption.

Accuracy with respect to selection percentage of trustworthy servers

The accuracy of the model is to represent the percentage selections, i.e., Number of times when it successfully selects trusted sensors out of total number of transactions. Figure 5 compares the accuracy of BTRM, LFTM, Power and Peer trust models. From the Fig. 5 we can conclude that when the percentage of malicious sensors is not high (less than 50 %) the difference between their accuracy is not significant and they can all achieve an accuracy of 97 %. However, when the situation is unsecured (percentage of malicious sensor is getting higher) then the Fig. 5 have shown that BTRM-WSN remains resilient to a high percentage of malicious servers when this percentage is less than or equal to 80 % (which is, however, a high percentage). But BTRM performance gets worse when there are 90 % or more

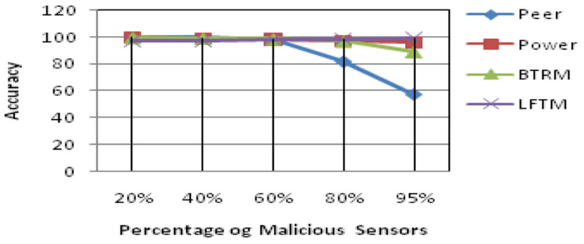


Fig. 5 Accuracy in searching for trustworthy sensors

Table 4 Accuracy

	20 %	40 %	60 %	80 %	95 %
LFTM	97.12	97.04	98.22	98.53	99.12
Peer	99.75	99.35	97.74	81.79	57.23
Power	99.74	99.24	98.44	97.4	96.39
BTRM	99.7	99.2	98.57	96.89	88.92

malicious servers in the WSN, and the size of the WSN grows. Table 4 shows the accuracy of the BTRM, Peer Trust system, Power Trust system and LFTM with different percentage of malicious nodes in network. LFTM performance is better even the number of malicious node is more than 80 %. Therefore, in smaller WSNs LFTM would work properly despite of a high percentage of malicious servers. The results show that for the heterogeneous case, it actually obtains better results as the number of untrustworthy serves increases. The reason is that the ants spread a given total amount of pheromone and that when the number of good servers is small, the paths to these are more strongly selected. In a way, the fewer the number of good servers is, the easier is for them to shine or excel. The results for the heterogeneous experiment, which is harder than the homogeneous one, are still highly successful regarding locating trustworthy servers over 90 % of the cases in the worst case and over 95 % when there are a few good servers.

Average path length leading to trustworthy servers

Path length consists of measuring the length (number of hops) of those paths found by TRM leading to trustworthy servers. That is, when the model fails and selects an untrustworthy server, that path is discarded and not taken into account. It is assumed that less average path length indicates a better performance in efficiency and easiness in searching for trustworthy sensors of a trust and reputation system as shown in Table 5. Figure 6 shows that BTRM and LFTM perform better and Peer and Power trust has the worst performance. LFTM trust model performs better when number of malicious node is above 90 %. With respect to path length a similar effect with the selection of trustworthy servers happens. We can consider that both models provide an easy and efficient approach in searching trustworthy sensors.

Energy consumption

How to efficiently reduce energy consumption is a major issue for WSN researchers. Figure 7 compares the 4 major Trust models Peer, LFTM, Power Trust and BTRM in terms of overall network energy consumption. The overall energy consumption means the energy consumed by client nodes while sending the request message, server nodes sending response message, energy consumed by malicious nodes, relay nodes and the

Table 5 Average path length

	20 %	40 %	60 %	80 %	95 %
LFTM	3.12	2.52	2.28	2.22	2.21
Peer	6.46	6.84	6.68	6.66	6.61
Power	6.7	6.87	6.73	6.8	6.85
BTRM	2.24	2.33	2.42	3.03	6.29

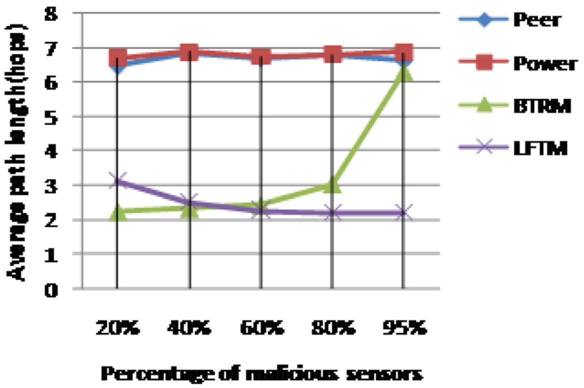


Fig. 6 Average path length leading to trustworthy sensors

energy to execute the trustworthy sensor searching process of a certain trust reputation system. Table 6 shows the detailed energy consumption by different trust systems. As shown in Fig. 7 the power Trust System consumed more energy and LFTM consumed lowest energy.

Performance evaluation of proposed LDAT

In RDAT, functional reputation is used to compute trust and reputation values based on three specific parameters, i.e., sensing, aggregating, and routing. These parameters were considered to evaluate the trustworthiness of the nodes. The role of the aggregator node is dynamically changing in the proposed protocol for security and energy saving. In our proposed protocol, there is no extra overhead for key distribution and sharing of keys for encryption and decryption.

In iRTEDA, Reputation and trust of sensor nodes are evaluated by other nodes of the same cluster. A watchdog mechanism is used to monitor the behavior of the neighbor nodes and actions are characterized as cooperative and non cooperative.

Sensor nodes sense the data, monitor the activities of other nodes, exchange the observations with the neighbouring nodes, calculate the trustworthiness of the nodes, and transmit the data and other information to the aggregator. Another issue is key distribution and sharing of keys for encryption/decryption of data between pairs of sensor nodes.

LDAT does not require any watchdog mechanism to evaluate the trustworthiness of the sensor nodes. No extra routing protocol is required in case of the proposed protocol and reliable path selection is performed along with trust evaluation. Thus,

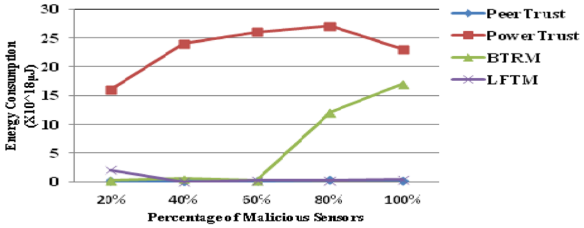


Fig. 7 Overall network energy consumption

Table 6 Energy consumption

	20 %	40 %	60 %	80 %	95 %
LFTM	2.1*10^17.0	0.01*10^17.0	0.24*10^17.0	0.22*10^17.0	0.36*10^17.0
Peer	0.15*10^17.0	0.10*10^17.0	0.17*10^17.0	0.28*10^17.0	0.26*10^17.0
Power	16*10^17.0	24*10^17.0	29*10^17.0	24*10^17.0	22*10^17.0
BTRM	0.21*10^17.0	0.58*10^17.0	0.22*10^17.0	12*10^17.0	21*10^17.0

there is no extra computation overhead. In RDAT there is no recovery mechanism, but in case of LDAT and iRTEDA, recovery mechanism exists. After identification of malicious nodes on the basis of malicious activity, a recovery mechanism is used to isolate the malicious nodes from the network. Under certain level security may be sacrificed to achieve lower energy consumption since sensor nodes are so limited in power resources [22].

Energy consumption

While implementing RDAT, four components of RDAT (a) monitoring component to evaluate the sensing, routing and aggregation misbehaviours (b) reputation component that computes the trust levels of sensor nodes (c) RDA for performing reliable Data Aggregation (d) RMDT for multipath data transmission to base station are considered. In LDAT protocol there is no separate method for selecting reliable path and with the addition of data aggregator node in the network the overhead and network traffic also gets reduced.

Figure 8 shows that protocol RDAT and iRTEDA has more energy consumption of sensor node due to more number of message transfers. The RDA and multipath data transmission components of protocol RDAT incurs higher number of message exchanges due to which more energy is consumed. Figure 8 shows the energy consumed by the RDAT and LDAT protocol. The Fig. 8 shows that RDAT and iRTEDA has more energy consumption as compared to LDAT.

Evaluation of accuracy

In RDAT sensing misbehaviours are detected via direct comparison of sensor node sensing values. In RDAT if 30 % of sensor nodes are compromised the accuracy of the aggregated data goes to 78 %., in case of iRTEDA accuracy is 82 %. In LDAT

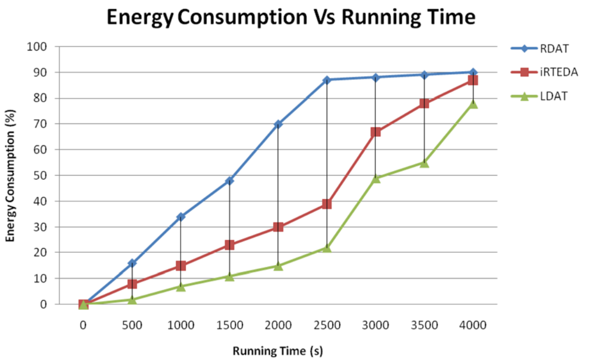


Fig. 8 Average energy consumption

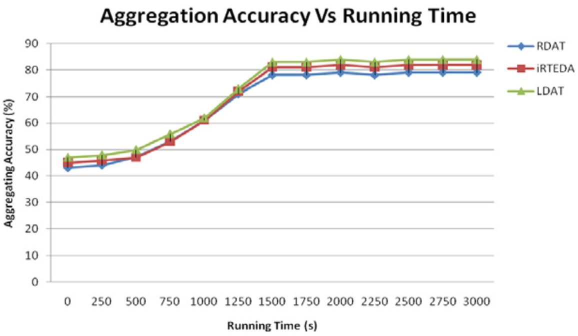


Fig. 9 Accuracy under the presence of 30 % compromised sensor node

the accuracy is above 84 %. In our simulation scenarios, we are considering both the presence of compromised sensor nodes and data aggregator nodes. Figure 9 shows the accuracy of both LDAT, RDAT and iRTEDA under the presence of 30 % compromised nodes.

Packet delivery ratio

Packet Delivery Ratio (PDR) is defined as the ratio of the total amount of packets reached at receiver and the number of packets sent by the source. If the number of malicious nodes increases, PDR also decreases gradually. Consecutely, throughput decreases as the amount of malevolent nodes increases.

$$PDR = number\ of\ packets\ reached\ at\ the\ destination/Number\ of\ packets\ sent\ by\ the\ source \quad (9)$$

The Packet delivery ratio of LDAT, RDAT and iRTEDA algorithms is shown in the Fig. 10. At this point we compare the proposed LDAT algorithm with RDAT and iRTEDA. The delivery ratio decreases with the increase in the fraction of misbehaving nodes consistently. The fraction of misbehaving nodes varied between (0–40) percent. When it increases the RDAT shows 0.65, iRTEDA shows .67 and LDAT shows 0.69.

Conclusion

In our proposed approach, Data Aggregation process of the network is made trustworthy by applying modified LFTM. *m*LFTM has applied linguistic fuzzy logic and

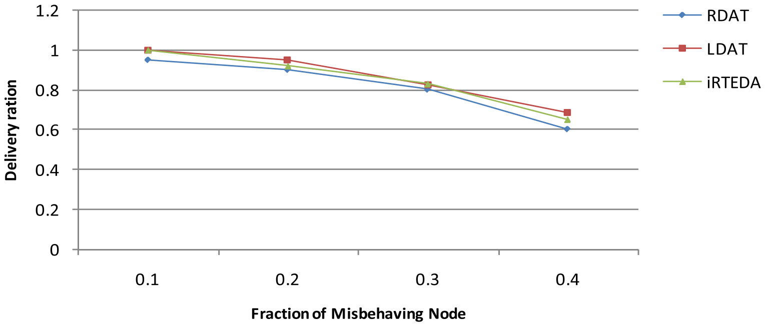


Fig. 10 Packet Delivery Ratio based Vs. percentage of fractions of misbehaving nodes

fuzzy sets to a previous bio-inspired trust and reputation model for WSNs. The reason of using *m*LFTM for securing Data Aggregation is that it enhances the interpretability of the model and makes it more users friendly. In our simulation result it is shown that our proposed approach, i.e., LDAT has less energy consumption, good packet delivery ratio and more accuracy than existing RDAT and iRTEDA under different scenarios. The performance of the protocol is evaluated in the presence of compromised data aggregator node and if the performance of the protocol decreased, changes are proposed to make the aggregation process more accurate and energy efficient.

Competing interests

The authors declare that they have no competing interests.

Received: 20 November 2014 Accepted: 22 March 2016
Published online: 19 April 2016

References

1. Liu C, Liu Y, Zhang ZJ (2013) Improved reliable trust-based and energy-efficient data aggregation for wireless sensor networks. Int J Distrib Sens Netw 2013(652495):1–11
2. Srinivasan A, Teitelbaum J, Liang H, Wu J, Cardei M. In: A. Boukerche (Ed.), "Reputation and Trust-Based Systems for Ad Hoc and Sensor Networks", Algorithms and Protocols for Wireless Ad Hoc and Sensor Networks, Wiley and Sons, 2008.
3. Misra S, Vaish A (2009) Reputation-based role assignment for role-based access control in wireless sensor networks. J Comp Commun 34(3):281
4. Suat O (2008) Functional reputation based reliable data aggregation and transmission for wireless sensor networks. Comp Commun 31:3941–3953
5. Marmol FG, Perez GM (2010) Towards Pre-standardization of trust and reputation models for distributed and heterogeneous systems. Comput Stand Inter 32(4):185–196
6. Marti S, Garcia-Molina H (2006) Taxonomy of trust: categorizing P2P reputation systems. Comput Netw 50(4):472–484
7. Sohraby K, Minoli D, Znati T (1991) Wireless sensor networks-technology, protocol and applications, Secondth edn
8. Mármol FG, Perez GM (2011) Providing trust in wireless sensor networks using a Bio-inspired technique. Telecommun Syst 46(2):163–180
9. Kumar EGP, Titus I, Thekkekara SI (2012) A comprehensive overview on application of trust and reputation in wireless sensor network. Proc Eng 38(2012):2903–2912
10. Xiong L, Liu L (2004) PeerTrust: supporting reputation-based trust for peer-to-peer electronic communities. IEEE Trans Knowl Data Eng 16(7):843–857
11. Zhou R, Hwang K (2007) Power trust: a robust and scalable reputation system for trusted peer-to-peer computing parallel and distributed systems. IEEE Trans Parallel Distrib Syst 18(4):460–473
12. Marzi H, Li M (2013) An enhanced Bio-inspired trust and reputation model for wireless sensor network. Proc Comp Sci 19(2013):1159–1166
13. Mármol FG, Marín-Blázquez JG, Perez GM (2012) LFTM, linguistic fuzzy trust mechanism for distributed networks. Concurr Comp Pract Exp 24(17):2007–2027
14. Chan H, Perrig A, Przydatek B, Song D (2007) SIA: secure information aggregation in sensor networks. J Comput Secur 15:69–102
15. Ganeriwal S, Srivastava MB (2004) Reputation-based framework for high integrity sensor networks. Proc. of the 2nd ACM workshop on Security of ad hoc and sensor networks, Washington DC, pp 66–77
16. Zhang W, Das SK, Liu Y (2006) A trust based framework for secure data aggregation in wireless sensor networks, Sensor and Ad Hoc communications and networks. SECON'06. 2006 3rd annual IEEE communications society. IEEE, Reston, pp 60–69
17. Wu K, Dreef D, Sun B, Xiao Y (2007) Secure data aggregation without persistent cryptographic operations in wireless sensor networks. Ad Hoc Netw 5(1):100–111
18. Bhavna AM, Padha DA (2010) Trust based secure data aggregation protocol in wireless sensor networks. IUP J Inform Technol 6(3):7–22
19. Nagarathna K, Kiran YB, Mallapur JD, Hiremath S (2012) Trust based secured routing in wireless multimedia sensor networks, Fourth International conference on computational intelligence, communication systems and networks (CICSyN'12)., pp 53–58
20. Dorigo M, Birattari M, Stutzle T (2006) Ant colony optimization. IEEE Comput Intell Mag 1(4):28–39
21. Marmol FG, Perez GM (2009) TRMSim-WSN, trust and reputation models simulator for wireless sensor networks, communications, IEEE international conference on. IEEE, Dresden, pp 1–5
22. Akyildiz IF, Su W, Sankarasubramaniam Y, Cayirci E (2002) Wireless sensor networks: a survey. IEEE Comp 38(4):393–422