Secure Clustering Algorithm for Wireless Sensor Networks in the Perspective of Authentication

Etuk, Enobong E. #1, S. Magesh, Ph.D. #2

Department Of Information Technology, Faculty of Engineering, SRM University, kattankulathur, Tamil Nadu-India

Abstract — Authentication happens to be a fundamental security service in cluster based wireless sensor networks. This project will focus on designing a lightweight authentication scheme that will aim at achieving authentication of cluster members with cloud enabled cluster heads. This model also presents a hierarchical and on-demand cluster maintaining method in the clustered wireless sensor networks, ensuring that the network can identify and get rid of malicious nodes, achieve cluster-formation, self-management and self-sustenance based on the random/parallel cluster header selection algorithm and a pre-distributed key scheme. This project will address secure mechanism on nodes joining in or quitting from network and a secure clustering algorithm based on the ant colony approach is proposed to make sure that mobile nodes joining in and quitting from network safely is based on authentication and secret telecommunications. An integrated solution is designed and developed to ensure efficient management, energy saving and information security for wireless sensor network, simulation results will show that the proposed model is promising in terms of networking security, efficiency and adaptability.

Keywords — wireless sensor networks, Authentication, cloud, node joining in and quitting from network; security; clustering; mechanism.

I. INTRODUCTION

A wireless sensor network (WSN) is a distributed autonomous sensor which has been designed to monitor physical or environmental conditions, such as temperature, sound and pressure and to pass the sensed data through the network to a main location such as a base station or cloud. The development of wireless sensor networks was motivated by military applications such as battlefield surveillance; today these networks are used in many industrial and consumer applications, such as industrial process monitoring and control, machine health monitoring, and so on. The networks are made up of hundreds or thousands of self-organizing, low-power and low-cost wireless nodes which are randomly deployed to gather enough information an object or event, used in monitoring the environment or tracking certain target to support decision[1]. Important requirements for these networks include prolonged network lifetime, scalability, and information security. sensor nodes clustering provides an effective technique for achieving these goals, and the clustered hierarchy provides an efficient approach to ensure information security, efficient-energy management and adaptability of the WSNs.

The WSN made up of “nodes” – ranging from a few to several hundreds or even thousands, where each node is connected to one (or several) sensors as may be required. Each sensor network node has several parts such as: a radio transceiver with an internal antenna or connection to an external antenna, a microcontroller, an electronic circuit for interfacing with the sensors and an energy source such as a battery or an embedded form of energy harvesting. The size of a sensor node might vary from that of a shoebox down to the size of a grain of dust. Its cost is similarly variable, ranging from a few to hundreds of dollars, depending on the complexity of the individual sensor nodes. The Size and the cost constraints on sensor nodes result in corresponding constraints on resources such as energy, memory, computational speed and communications bandwidth. The topology of the WSNs can vary from a simple star network to an advanced multi-hop wireless mesh network. The propagation technique between the hops of the network may be routing or flooding.

A. OBJECTIVE

This project aims at designing a secure clustering algorithm for wireless sensor networks based on THE ANT COLONY APPROACH. The network will be organized in a way that each cluster will have a cloud enabled cluster head that will perform data fusion and aggregation on behalf of all other cluster members. Each cluster is autonomous (i.e. self-organizing and self-sufficient) but cluster members are in turn mobile nodes. In the event of low energy level in any cluster member or a cluster head requires more cluster members for efficient information gathering, the cloud enabled cluster head will broadcast a request to neighbouring cluster specifying its needs. A node will be assigned to the cluster in need by the neighbouring cluster head. The node will present its ID for authentication while all other authentication credentials will be forwarded by its cluster head. This is to ensure that nodes joining –in or quitting from the network are properly authenticated to ensure integrity.
of the information gathered. This wireless sensor network may be used for habitat monitoring.

Some GOALS this project aims to achieve include:

- Introduce secure doings into clustering process before all of nodes begin to collect and transmit data.
- Identify malicious nodes in the clustering phase and excluded them from data transmission.
- Malicious nodes will be prevented from joining the Cluster and valid nodes will be authenticated via a lightweight authentication scheme.
- Prevent hostile nodes gaining access to the cluster.

B. SCOPE

There are several key limitations in WSNs that the proposed secure clustering algorithm aims to address such as:

- **Limited Energy**: Unlike wired designs, wireless sensor nodes are “off-grid”, meaning that they have limited energy storage and the efficient use of this energy will be vital in determining the range of suitable applications for these networks. The limited energy in sensor nodes will be considered as proper clustering can reduce the overall energy usage in a network.

- **Network Lifetime**: The energy limitation on nodes results in a limited network lifetime for nodes in a network. Proper clustering should attempt to reduce the energy usage, and thereby increase network lifetime.

- **Limited Abilities**: The small physical size and small amount of stored energy in a sensor node limits many of the abilities of nodes in terms of processing and communication abilities. A good clustering algorithm should make use of shared resources within an organizational structure, while taking into account the limitation on individual node abilities [15].

- **Application Dependency**: Often a given application will heavily rely on cluster organization. When designing a clustering algorithm, application robustness must be considered as a good clustering algorithm should be able to adapt to a variety of application requirements.

II LITERATURE REVIEW

To the best of our knowledge there are several typical clustering algorithms emerged such as LEACH, HEED, SLEACH, SEFA AND SPLEA, SHEER algorithms.

**LEACH**[2] is a self-adaptive and periodical clustering algorithm. LEACH organizes nodes into clusters with one node from each cluster serving as a cluster-head. Nodes first send sensor readings to their cluster-head, and the cluster-head aggregates or compresses the data from all its “children” for transmission to a base station. LEACH uses randomized rotation of nodes required to be cluster-heads to evenly distribute energy consumption over all nodes in the network. LEACH operation is broken into rounds, with each round having a set-up phase and a steady state phase. Its obvious shortcoming is the election of cluster head, it doesn’t take the remaining energy of node into account and the distribution of cluster head is not uniform, this will weaken the coverage and lifetime of network. Also, since nodes choose a cluster-head based on received signal strength, an adversary can disable the entire network by using the HELLO flood attack to send a powerful advertisement to all nodes in the network. Due to the large signal strength of the advertisement, every node is likely to choose the adversary as its cluster head. The adversary can selectively forward those data that actually reach her, while the rest of the network is effectively disabled.

**HEED**[3] proposes a new energy-efficient approach for clustering nodes in ad hoc sensor networks. It presents a protocol that periodically selects cluster heads according to a hybrid of their residual energy and a secondary parameter, such as node proximity to its neighbors or node degree. HEED does not make any assumptions about the distribution or density of nodes, or about node capabilities. The clustering process does not depend on the network topology or size. The protocol incurs low overhead in terms of processing cycles and messages exchanged. It also achieves fairly uniform cluster head distribution across the network. A careful selection of the secondary clustering parameter can balance load among cluster heads. It is noted that nearly all of the existed clustering algorithms including above two ones is dependent on the assumption that the surrounding is non-hostile, they cannot identify and refuse the malicious nodes for lack of secure mechanism and authentication, there is enormous risk hidden and this will result in a sharp increase in cost and difficulty to realize security goals at higher layers, for example, routing or data aggregating.

**SEFA AND SPLEA** Sudarshan Vasudevan [4] and et al consider the problem of secure leader election and propose two cheat-proof election algorithms : Secure Extrema Finding Algorithm (SEFA) and Secure Preference-based Leader Election Algorithm (SPLEA). Both algorithms assume a synchronous distributed system in which the various rounds of election proceed in a lock-step fashion. SEFA assumes that all elector-nodes share a single common evaluation function that returns the same value at any elector-node when applied to a given candidate-node. When elector-nodes can have different preferences for a candidate node, the scenario becomes more complicated. SPLEA deals with this case. Here,
individual utility functions at each elector-node determine an elector-node’s preference for a given candidate-node. But the assumption is rigorous and invalid when any malicious node emerges. **SLEACH** Wang Xiao-yun[5] and al proposed the SLEACH algorithm as a security extension of LEACH one to withstand adversaries or the compromised nodes, the sink node authenticates all candidate cluster-heads with enormous added traffic and secure key requirement, therefore, it is not a perfect solution to security-characteristic wireless sensor network. **SHEER algorithm**[7] improves the energy-efficient consumption and lifetime of network by a random broadcasting system and the three-tier model of clustering.

Algorithm. **ANT COLONY OPTIMIZATION** algorithms have been applied to many combinatorial optimization problems and a lot of derived methods have been adapted to dynamic problems in real variables, stochastic problems, multi-targets and parallel implementations. It has also been used to produce near-optimal solutions to the travelling salesman problem. They have an advantage over simulated annealing and genetic algorithm approaches of similar problems when the graph may change dynamically; the ant colony algorithm can be run continuously and adapt to changes in real time. This is of interest in network routing and urban transportation systems.

The first ACO algorithm was called the Ant system and it was aimed to solve the travelling salesman problem, in which the goal is to find the shortest round-trip to link a series of cities. The general algorithm is relatively simple and based on a set of ants, each making one of the possible round-trips along the cities. At each stage, the ant chooses to move from one city to another according to some rules:

- It must visit each city exactly once;
- A distant city has less chance of being chosen (the visibility);
- The more intense the pheromone trail laid out on an edge between two cities, the greater the probability that that edge will be chosen;
- Having completed its journey, the ant deposits more pheromones on all edges it traversed, if the journey is short;
- After each iteration, trails of pheromones evaporate.

**III DESIGN PHILOSOPHY**

Wireless Sensor Networks present vast challenges in terms of implementation. Design goals targeted in traditional networking provide little more than a basis for the design in wireless sensor networks [16], [17], [18]. Clustering algorithms play a vital role in achieving the targeted design goals for a given implementation. There are several key attributes that designers must carefully consider, which are of particular importance in wireless sensor networks.

1) **Cost of Clustering**: Although clustering plays a vital role in organizing sensor network topology, there are often many resources such as communication and processing tasks needed in the creation and maintenance of the clustering topology.

2) **Selection of Cluster heads and Clusters**: The clustering concept offers tremendous benefits for wireless sensor networks. However when designing for a particular application, designers must carefully examine the formation of clusters in the network. Depending on the application, certain requirements for the number of nodes in a cluster or its physical size may play an important role in its operation. This prerequisite may have an impact on how cluster heads are selected in this application.

3) **Real-Time Operation**: Useful lifetime of data is also a fundamental criterion in designing Wireless Sensor Networks. In applications such as habitat monitoring [19], [20], simply receiving data is sufficient for analysis, meaning delay is not an important issue but When we consider military tracking [21], the issue of real-time data acquisition becomes much more vital. Also cluster recovery mechanisms must also be taken into account.

4) **Synchronization**: One of the primary limitations in Wireless Sensor Networks is the limited energy capacity of nodes. Slotted transmission schemes (such as TDMA), allow nodes to regularly schedule sleep intervals to minimize energy used. Such schemes require synchronization mechanisms to setup and maintain the transmission schedule.

5) **Data Aggregation**: One major advantage of wireless sensor networks is the ability for data aggregation to occur in the network. In a densely populated network there are often multiple nodes sensing similar information. Data aggregation allows the differentiation between sensed data and useful data. Network processing makes this process possible and now it is fundamental in many sensor network schemes [22], as the power required for processing tasks is substantially less than communication tasks. As such, the amount of data transferred in network should be minimized.

6) **Repair Mechanisms**: Due to the nature of Wireless Sensor Networks, they are often prone to node mobility, node death and interference. All of these situations can result in link failure. When looking at clustering schemes, it is important to look at the mechanisms in place for link recovery and reliable data communication.
7) **Quality of Service (QoS):** From an overall network standpoint, we can look at QoS requirements in Wireless Sensor Networks. Many of these requirements are application dependent (such as acceptable delay and packet loss tolerance), and as such, it is important to look at these metrics when choosing a clustering scheme. Implementations can vary widely in terms of these metrics, and as a result, the design process should consider these aspects.

**A. PROPOSED SOLUTION**

Secure clustering algorithm is an algorithm that usually has one or both of the following goals in solving a problem:

- Finding an algorithm with reasonable run-time (time needed to set up clusters is affordable).
- With finding the optimal solution. This means that a heuristic algorithm leads to reasonable performance and is not based on particular metrics.

The destroying behavior of malicious nodes in wireless sensor networks usually includes:
- forging data
- refusing to transmit data
- largely and repeatedly transmitting real or false data to exhaust the energy of other nodes
- Routes misguidance.

The thought of a secure clustering algorithm (SCA) in wireless sensor networks is to introduce secure doings into clustering process before all of nodes begin to collect and transmit data. Any malicious node will be identified in the clustering phase and they will be excluded usual data transmission later, also malicious nodes joining will be prevented from joining the Cluster via a lightweight authentication scheme. This is an original solution with more advantages to prevent hostile nodes gaining access to the cluster. The SCA algorithm is separated into secure cluster-head election, secure node selection, secure data transmission and secure cluster maintenance and finally eventually deals with node cheating and false information problems.

**B. Secure Cluster-Head Election**

On accomplishing deployment of nodes in WSNs, any non-clustered node may randomly initiate to set up cluster, the neighboring nodes securely and dynamically elect cluster-head and form cluster according to the shared secret knowledge, the residual energy of nodes and the clustered status in quo. Those nodes entitled to initiating to set up cluster are:
- Any of new nodes deployed
- A clustered nodes that has lost touch with its cluster-head
- Any cluster node which gives up as a cluster-head because of heavy energy consumption.
In fact, nodes of this kind are non-clustered now.

**C. Secure Node Selection**

Any non-clustered node may declare itself as a temporary cluster-head and advertises its ID and clustering-order message which is encrypted via the pre-mounted main key to its neighboring nodes. All nodes which receives the message must will send back an answer-message to the temporary cluster-head, indicating whether the sender is cluster-head or temporary cluster-head or clustered node or not, and its ID and residual power. Any of neighboring nodes which receives the clustering-order message checks the received ID to ensure the message is not from a malicious. It does this by checking in a blacklist of IDS; if it has drawn a conclusion that the message is from a certain hostile node, then it will not respond to the request. After passing the secure check, if the node which received the clustering-order message is not a clustered yet, then it should send back a message including own ID and residual energy information to the temporary cluster-head node. If it is clustered, it sends back a message including its own ID and cluster-head tag. The temporary cluster-head node makes an ID check on the feedback information it gets. Only the non-malicious node can receive and read any other messages that come from the temporary cluster-head.

Finally, the cluster-head node advertises the encrypted clustering message to its neighboring nodes, whose cluster-head tag and ID are encapsulated in the message. The neighboring non-clustered nodes that receive the message will become a cluster-head node of the cluster and send back their IDs and residual energy information to the sending cluster-head node, others will keep silent. If those neighboring non-clustered nodes receive several invitations of clustering at the same time, each of them will pick up the closest cluster-head and join its cluster based on the largest received signal strength of the advertisement from the corresponding cluster-head.

The above clustering process will be repeated as long as any node issues the clustering request, till each of the non-malicious nodes becomes a member of certain cluster.

**D. Secure Data Transmission**

After setting up the clusters, the cluster-head of any cluster will create a TDMA schedule [6] telling each node when it can transmit data based on the number of nodes in the cluster.

The node always sends encrypted data to the cluster head during its allocated transmission time; at this time, all other nodes are OFF. The cluster-head keeps its receiver on to receive all the data from the nodes in the cluster and confirms the authenticity of all received data based on the shared master key. When all data has been received, the cluster head performs the data aggregation and sends the composite signal to the control station under the control of the secure protocol via the planned routes. Any data transmission will be strictly authenticated via the shared key to distinguish...
the identities of the corresponding partners and the origin of the received message, the two-way authentication will be executed via a CHALLENGE-AND-RESPONSE process.

E. Secure Cluster Maintenance

After wireless sensor networks run period of time, the residual energy of nodes will be different for each node due to their diverse roles and different energy consumption. The cluster head node, for example uses more energy than the other nodes due to its role and the task it performs. Therefore, some nodes may die because of energy exhaustion, while some new nodes full of power are distributed to join the network. To prolong the lifetime of network and to ensure operational effectiveness and convenience, a cluster-head may request new nodes from neighboring clusters thus, a new and stable clustering process is necessary for the authentication of new nodes by the aforementioned method; this is the secure cluster maintenance.

The cluster head maintenance is likely to take place at any time after cluster setup phase, which will be executed on request, not periodically, and the secure cluster maintenance happens locally in a certain scope of network in which non-clustered nodes are present or a neighboring cluster-head has nodes to spare.

The cluster head node will periodically compare its residual energy with that of other nodes in the cluster, if its residual energy is lesser than the mean value of the others, the cluster head node will give up as the cluster head and puts up a clustered node which holds most residual energy as new cluster head and safely advertises the message of cluster head shift to the other nodes in the cluster.

IV. SOLUTION ARCHITECTURE

A. Sensor Network Architecture

One of the major problems in sensor networks is how to create an organizational structure amongst these nodes [15]. Since the fundamental advantage of WSNs is the ability to deploy them in an ad hoc manner, as it is not feasible to organize these nodes into groups pre-deployment. For this reason, there has been a large amount of research into ways of creating these organizational structures (or clusters) [13], [14], [15].

- **Sensor Node** [23]: A sensor node is the core component of a WSN. Sensor nodes can take on multiple roles in a network, such as simple sensing; data storage; routing; and data processing.

- **Clusters**: Clusters are the organizational unit for WSNs. The dense nature of these networks requires the need for them to be broken down into clusters to simplify tasks such as communication.

- **Cluster heads**: Cluster heads are the organization leader of a cluster. They often are required to organize activities in the cluster. These tasks include but are not limited to data-aggregation and organizing the communication schedule of a cluster.

- **Base Station**: The base station is at the upper level of the hierarchical WSN. It provides the communication link between the sensor network and the end-user.

- **End User**: The data in a sensor network can be used for a wide-range of applications. [8] Therefore, a particular application may make use of the network data over the internet, using a PDA, or even a desktop computer. In a queried sensor network (where the required data is gathered from a query sent through the network). This query is generated by the end user. The clustering phenomenon as we can see, plays an important role in not just organization of the network, but can dramatically affect network performance.

![Fig. 1 A sample wireless sensor network cluster comprising of three clusters each having a cluster head and sensor nodes and all connecting to a base station.](image)

B. Authentication Process

The authentication of new nodes to join any cluster is based on a CHALLENGE-AND-RESPONSE process.

When a node is not able to sense a parameter for want of energy, the cluster-head (CH1) sends a broadcast to neighboring cluster-heads to allocate some of its mobile nodes to its cluster. A neighboring cluster-head (CH2) allocates a node to CH1 cluster. When a mobile node tries to join the cluster, CH2 sends a challenge RAND to CH1 and to the new node (N1). CH1 computes a response (RS) using RAND, CH1 tag and an authentication algorithm (a lightweight authentication algorithm). N1 also computes its own response (RN) using its ID, RAND, CH1 tag and same authentication algorithm and sends RN, encrypted with a key derived from the shared master key(between Cluster-heads) to CH1. CH1 compares the received RN with its own computed R. If they match, then N1 is considered non-malicious and is authenticated. CH1 then broadcast N1 information to other nodes in the cluster and to the control station for update.
NOTATIONS:
CH1 - cluster-head of requesting cluster
CH2 - cluster head of sending cluster
N1 - mobile sensor node from sending cluster
R - CH1 response
RS - N1 response
RAND - random number for challenge response computation
Kcl - cluster key
Ks - Sensor node key

Fig.2. The authentication process when a node is sent from a cluster to another

C. Key Management Scheme

Symmetric key mechanism is used by each node; therefore each node is expected to store the keys it shares with its higher levels of the network hierarchy. Since the sensors are memory constrained and are susceptible to attacks by the adversary, they should be assigned the minimum number of keys which in turn saves memory. If a node is compromised it will reduce the impact of the damage since less number of keys would be revealed to the adversary. As soon as the sensor nodes are deployed, all the sensor nodes and clusters send their ID to the base station. The key management scheme consists of three keys:

- Kg (group key) - Generated by base station, and pre-deployed to all sensor nodes in the entire network. Node uses this key to encrypt the data.
- Kcl (cluster key) - Generated by base station and deployed to all cluster heads in the entire network. Base station and nodes from the cluster head use this key to decrypt the data.
- Ks (sensor key) - Generated by the base station and deployed to all sensors in the entire network. The base station uses this key to encrypt the data to send to the cluster head.

when base station sends a message to the cluster head, it constructs the message as follows:

\{ Kg, Kcl, MAC, ID, TS, N, Message \}

Base station encrypts its own ID and a current time stamp TS. It generates a random number N and Kcl for cluster header.

Cluster head checks the ID from the packet to ensure it matches the ID it holds and verifies the authentication and integrity of the packet through MAC, otherwise packet is dropped by the cluster head. The base station builds the message using the fields below:

Eks \{ Kg, Kcl, MAC, ID, TS, N, Message \}

Base station encrypts the message and broadcast the data. When the cluster head receives the messages, it decrypts it by using Ks. Cluster head aggregates the messages received from its nodes and forwards it to the next level cluster head or if the cluster head is one hop closer to sensor node. Receiving cluster head checks its routing table and constructs the following packets to be sent to the next level cluster head. The cluster head adds its own ID and rebuilds the packet at the following:

\{ IDcl, \{ Kg, Kcl, MAC, IDcl, TS, N, Message \} \}

Here the ID is the ID of the receiving cluster head which wraps the message and sends to the next hop cluster head or to the lowest cluster head. Next hop cluster head receives the packet and checks its own ID, if the ID in the packet is the same as its holds, it updates the ID for the next hop and broadcast it, or else the packet is discarded. When cluster head sends the message to the sensor nodes, it constructs the messages as follows:
Finally, the sensor node gets Kg for use to encrypt the message in order to send to the cluster head. All the keys are generated by the base station. Base station also verifies MAC and data and also process and aggregate data. Cluster heads aggregate data coming from the sensor nodes and then sends to the base station.

D. Analysis of Secure Clustering Algorithm

The main features of this algorithm are: (1) A node is selected as a cluster head based on the election rules and it residual energy. A rapid and efficient hierarchical topology is setup and routes established among all cluster head nodes in a large-scale network. (2) The network can adapt to change of topology. Compromised nodes are removed from the network and new nodes are added securely. (3) Residual energy of nodes are monitored to ensure survivability of the networks. (4) The algorithm provides good connectivity between nodes and base station. (5) The two-way authentication between sender and receiver helps in detection of compromised nodes during the clustering phase and the nodes are evicted from the clustering process ensuring security of routing process and authenticity of aggregated data.

The proposed algorithm is evaluated by experiment and arrives in a promising result as regards networking security, efficiency and adaptability. The next phase would be to access the possibility of applying Ant colony optimization in depth in conjunction with the proposed authentication algorithm.

V. CONCLUSION

Based on existing concepts as regards security in wireless sensor networks, this algorithm (1) Proposes to provide secure authentication of nodes while joining or quitting a cluster thereby improving the security of the network and authenticity of aggregated data. Malicious nodes are gotten rid of using the proposed key management scheme which is based on nodes ID and pre-distributed keys. (2) The clustering algorithm integrates efficiency of energy utilization, clustering and dynamic topology, which will in turn improve and extend the lifetime and robustness of the network. This is a novel approach to provide authentication, conserve energy and provide secure clustering in wireless sensor networks. This algorithm may be extended to accommodate other relevant areas as regards the optimal operability of wireless sensor networks.

REFERENCES