Multi-stage secure clusterhead selection using discrete rule-set against unknown attacks in wireless sensor network

Tejashwini N.1, D. R. Shashi Kumar2, K. Satyanarayana Reddy3
1Visvesvaraya Technological University, India
2Department of Computer Science and Engineering, Cambridge Institute of Technology, India
3Department of Information Science and Engineering, Cambridge Institute of Technology, India

Article Info

ABSTRACT
Security is the rising concern of the wireless network as there are various forms of reconfigurable network that is arised from it. Wireless sensor network (WSN) is one such example that is found to be an integral part of cyber-physical system in upcoming times. After reviewing the existing system, it can be seen that there are less dominant and robust solutions towards mitigating the threats of upcoming applications of WSN. Therefore, this paper introduces a simple and cost-effective modelling of a security system that offers security by ensuring secure selection of clusterhead during the data aggregation process in WSN. The proposed system also makes construct a rule-set in order to learn the nature of the communication iinn order to have a discrete knowledge about the intensity of adversaries. With an aid of simulation-based approach over MEMSIC nodes, the proposed system was proven to offer reduced energy consumption with good data delivery performance in contrast to existing approach.

Keywords:
Cryptography
Encryption
Energy
Security
Wireless sensor network

Corresponding Author:
Tejashwini N.,
Visvesvaraya Technological University,
Belagavi, Karnataka, India.
Email: tejashwini.n@gmail.com

1. INTRODUCTION
The usage of wireless network has almost substituted majority of the legacy wired product in past decade. Although, there are significant advance of the wireless networks, there are potential challenges too [1]. There is various research works towards investigating the solutions for identifying the challenges in different forms of wireless network e.g. wireless local area network (WLAN) [2, 3], wireless sensor network (WSN) [4, 5], mobile adhoc network (MANET) [6, 7], etc. Out of all forms of wireless network, WSN has got a significant future as it is an integral part of Internet-of-Things (IoT) as well as any cyber-physical system [6]. There are wide ranges of application supported by WSN that has significantly found its way to the commercial products. However, security factor is becoming a rising concern because of various reasons. The first reason of security concern about WSN is that it cannot execute typical cryptographic programs as it doesn’t have adequate resources to execute complex cryptographic programs [8]. A complex cryptographic algorithm e.g. RSA (Rivest Shamir Algorithm) is a strongest encryption technique; however, size of the key for RSA is too large to be consistently used in a resource constraint sensors. The second reason of security concern is that majority of the existing security approaches are only meant for addressing specific form of attacks. This increases the development cost and reduces its applicability as malicious program can change at any point of time. Hence, the applicability of the security solution diminishes in this regard. The third reason of security concern is that it is not feasible for identifying the attacks generated from cross-platform networks. This usually happens when two different networking technologies are connected together e.g. IoT is formed by the integration of sensors and cloud. Where it is one of the most challenging tasks to identify
and resist the attacks from different network domain. The fourth reason of security concern is that it is quite challenging to identify the presence of attacker for a given scenario of wireless communication. Hence, irrespective of presence of many encryption methods over wireless network, there is still a huge research gap as none of the existing security-based approaches offers resistance to the ultimate threat [9-13]. It can be also said that the existing forms of hardware and network architecture of sensor node is definitely not ready to make secure transmission over wireless network.

At present, security approaches are designed using cryptographic approaches mainly which dominantly uses secret keys that are required to be stored safely within the nodes and will be used again. It will mean that usage of key-based encryption is only good when it have multiple numbers of passes to generate the secret key using some complex mathematical function. It will also mean that there will be higher usage of storage and processing capability of the node that will significantly drain its energy anyway. Therefore, the most significant eligibility factor of the robust encryption technique is that it should be lightweight. By lightweight, it will mean that the algorithm should have less iterations, it shouldn’t store much intermediate state message to avoid overheads. Hence, it is required that a good security algorithm should have better communication capability as well as lesser resource consumption for ensuring better network lifetime. Therefore, the proposed system targets to address the above-mentioned problems associated with security problem in wireless networking taking the case study of wireless sensor network (WSN). The paper also presents a simple and lightweight encryption methodology that is proven to offer good energy saving and better data delivery performance. Section 2 discusses about algorithm implementation followed by discussion of result analysis in section 3. Finally, the conclusive remarks are provided in section 4.

There are various amount of research-based contribution in order to deal with security problems in WSN. The recent work carried out by Zhang et al. [14] have introduced a security approach towards safeguarding the communication over relay networks with more emphasis on physical layer security. The problem of falsified injection of data is addressed most recently by Yang et al. [15] using a data fusion approach for advanced application of cyber-physical system. Tian et al. [16] have presented a cross-layer approach for enhancing the spectrum efficiency over multi-hop network. Huang et al. [17] have presented a group-key based security approach in order to address the communication overhead problems in WSN integrated with software device networks. Al-Turjman et al. [18] have presented a key-agreement scheme using public key encryption considering mobile sink. Zhao et al. [19] have implemented composite key Predistribution for securing the vulnerable communications. The approach of rekeying process has been revised in the approach of Aissani et al. [20] to generate a unique key management system with less overhead. Cao et al. [21] have presented an identification scheme of random attacks in WSN for securing the relay networks. Implementation of asymmetric encryption was carried out by Chen et al. [22] using hardware-based approach. A unique authentication protocol was presented by Chu et al. [23] using similar hardware-based approach that uses simple exclusive-or operation for encryption.

Adoption of chaotic approach was carried out in the work of Gopalakrishnan and Bhagyaveni [24]. The authors used this scheme for selection of the secure links with better data delivery performance. Hong et al. [25] have adopted a game theory-based model for resisting any form of attacks related to the interference. Li et al. [26] have implemented a generation approach for symmetric key using group-based communication system as well as signal strength factor. Shin et al. [27] have presented a concept of multi-factor authentication system that is meant for securing the WSN that is jointly operated with 5G technologies. Nkwe et al. [28] have presented security over the physical layer using error factor and power aspect for formulating security policies. Shim et al. [29] have presented an authentication approach using identity-based signature policy. A unique attack identification framework is presented by Sun et al. [30] that make use of principal component analysis in order to minimize the dependencies of more number of detection features. Adoption of symmetric key practices was witnessed in the work of Oliveira et al. [31] where low energy devices are used reducing the overheads of the control message. The technique uses advanced encryption standard for performing security. Pintea et al. [32] have emphasized on the usage of the sensitive agents in order to resists denial of service attacks in WSN. Han et al. [33] have implemented the concept of identifying the level of intrusion in order to resist the sinkhole attack in WSN. Therefore, it can be seen that there are different forms of approaches in existing system for resisting security threats in wireless network. The brief outlining of the problems associated with the existing system is discussed in next section.

The significant research problems are as follows:

- The security approaches of existing system are more focused towards addressing solution against specific security threat.
- Evidence towards resource-efficiency about the existing security approaches are not proven in existing literatures.
- Study towards less computational complexity factor associated with cryptographic usage towards securing vulnerable wireless network is less prominent.
Availability of less benchmarked solution towards potential resistance against lethal attack is found less in existing literatures. Therefore, the problem statement of the proposed study can be stated as “Developing a potential security solution to resist unknown forms of attack in wireless network using cost effective resistive solution with far reaching effect”. The next section outlines the solution towards resisting these problems.

The implementation of the proposed technique is carried out using an analytical research methodology and the proposed system acts as an extension of our prior work [34]. The core aim of the proposed technique is to offer a significant protection against an unknown form of adversary in WSN using simple and yet robust form of security mechanism. The schematic diagram to represent the implementation of proposed system is highlighted in Figure 1.

![Schematic Diagram](attachment:image.png)

**Figure 1. Proposed schematic diagram**

Figure 1 highlights the implementation flow of the proposed system where it can be seen that the proposed system considers initially a typical preparation of the simulation area by considering the sensors, distance, adversaries, and energy factor. The proposed system also considers that adversary module is quite potential; however, it will never launch any attack when it joins a new network. Hence, there is no apriori information about the adversary. The proposed system also constructs a rule-set that is further classified in two forms i.e. primary Ruleset and secondary rule-set. These rule sets are meant to perform a specific task of selection of secure clusterhead. The primary rule-set is meant for ensuring selection of most eligible clusterhead followed by secondary Ruleset that is meant to perform security analysis of the primarily selected clusterhead. The outcome of the secondary rule-set is a finally selected secured clusterhead. This operation is followed by constructing the design in order to investigate the intensity of the attacker followed by applying this rule set in order to ensure proper selection. Finally, the adversary node is identified and is updated to all the neighboring as well as finally to complete network so that all the communication process generated to and from such adversary could be terminated. The core objective of the proposed methodology is to ensure that a lightweight security policy is implemented in order to ensure that a better supportability of public key encryption is available.

## 2. ALGORITHM IMPLEMENTATION

The prime purpose of the proposed algorithm is to offer a secure data aggregation by addressing the secure selection of cluster-head. The hypothesis of the proposed algorithm construction is carried out on the fact that a clusterhead bears significant and valuable information as compared to other member nodes. Therefore, it is more likely that clusterheads will be more targeted to be compromised as compared to member nodes. Hence, the complete algorithm constructions are carried out to protect all sorts of communication originating to and from the clusterhead in WSN. The algorithm takes the input of $S_{area}$ (Simulation area), $E_i$ (initial energy), $n$ (sensors), $p_{x,y}$ (position of sensors), $p_{CH}$ (probability of cluster-heads), and $n_{ad}$ (number of adversary) that after processing yields an outcome of $n_{CH}$ (secure CH). The steps included in the proposed algorithm are as follows:
Algorithm for Selection of robust CH for secure Data Aggregation

Input: $S_{area}$, $E_i$, $n$, $p_{x,y}$, $P_{max}$, $n_{ad}$

Output: $n_{CH}$

Start
1. $init$ $S_{area}$, $E_i$, $n$, $p_{x,y}$, $P_{max}$, $n_{ad}$
2. $S_{area}\rightarrow\{b_{x,y}, rand(n_{x,y})\}$
3. construct $\phi_i, \phi_s$ // $\phi$ is Ruleset, $P$, $S$-primary/secondary
4. For $i=1:n$ //fuzzy+Active+dynamich 
5. $V_d \rightarrow \sum(\sqrt{(x_i-x(i))^2+(y_i-y(i))^2}$
6. For $j=1:n$
7. $N_{non}\rightarrow E_i(n, p_{x,y}, R_{an})$
8. $vul\rightarrow E_i(n, p_{x,y})$
9. $vic\rightarrow E_i(p_{x,y}, N, b_{x,y})$
10. $PO_i \rightarrow \phi_i(E_{N_{non}}) \& S\rightarrow \phi_s(vul, vic, d_{ch})$
11. $n_{max}\rightarrow argmax $(GQ1)$:
12. $End$
13. $End$
$End$

The flows of the proposed algorithmic steps are illustrated as below:

2.1. Preparation of the simulation area

The algorithm initializes all the input parameters (Line-1) and performs deployment of all the sensor nodes in random order i.e. $rand(n_{x,y})$, whereas it can also change the position of the base-station $b_{x,y}$ at any part of the simulation area (Line-2). The algorithm also selects a certain number of adversary nodes; however, in order to assess the system, the algorithm is designed with no pre-defined information about the location of the adversary.

2.2. Construction of rule-set

The proposed system constructs two sequential rule set that is utilized for filtering the selection process of the secure clusterhead. The first rule-set $\phi_r$ is about primary selection while second rule-set $\phi_s$ is about secondary selection criterion of the clusterhead (Line-3). Construction of both the rule-set are carried out using trapezoidal membership function.

2.2.1. Primary rule-set $\phi_r$

The algorithm takes the energy and the number of the neighbor as an input parameter to the rule-set processor that offers the output of primary selected clusterhead. The construction of the rule-set for inputs are carried out considering lower, higher, and medium value of if, whereas inferencing of the output is carried out by multiple combination of it as highlighted in Figure 2. The logic of this rule-set is that if a sensor node bears lower residual energy and lower neighbor nodes, those nodes don’t suit to be become strong clusterhead and hence never come under the radar of adversary. Apart from security viewpoint, such weaker form of nodes will not serve the purpose of being a clusterhead owing to their incapability to offer robust data aggregation. On the other hand, if the residual energy of a node is quite high and it has good connectivity with other adjacent nodes, those nodes are considered as the potential clusterhead. It is also imperative that such clusterhead will likely to attract the attention of the adversary as if such nodes are compromised then dispersion of the malicious attack will be more. Hence, the primary rule-set end up selecting the most suitable cluster head that can assists in data aggregation process for longer duration with good residual energy.

2.2.2. Secondary rule-set $\phi_s$

This rule-set is meant for further filtering the selection process of the secure clusterhead as an extension of the primary rule-set. Figure 3 highlights that in this stage, the processor takes three different inputs i.e. vulnerability, vicinity, and distance. All these parameters are more-or-less linked with distance metric itself. Vulnerability is basically condition when a sensor node is in a distance within the sensing range of an adversary while vicinity is same distance with closest distance with the adversary node. On the other hand, distance is basically a Euclidean distance between two communicating clusterhead (in case of multi-hop communication only). According to this final rule-set, the safest clusterhead will be that node which has lower value of vulnerability, vicinity, as well as distance. It simply infers that only node that are located at the farthest end from the adversary are the secure node and is more likely to be selected as clusterhead.
2.3. Assessing intensity of adversary

The proposed study constructs the adversary scenario with a combination of i) active attack and ii) dynamic state of node. Although, the algorithm is equally capable for testifying different possible attack scenario and different discrete states of node; this scenario is selected for analysis as this is the most challenging scenario with respect to intensity of attack and its aftermath consequences. The algorithm considers all the nodes present in the simulation area (Line-3) and implements a simple formula of $\theta$ (Line-4) in order to render arbitrary orientation of the sensor nodes. This part of the algorithm implementation is all about checking the distance to investigate the presence of adversary. For this purpose, the algorithm computes the vulnerable distance $V_d$ (Line-5). It further performs more filtration of the distance on the basis of vulnerability assessment using simple distance-based logic. For all the sensor nodes (Line-6), the proposed algorithm applies multiple explicit functions in order to perform this assessment. The algorithm first implements a function $f_1(x)$ that takes the input arguments of number of nodes ($n$), position of nodes ($p_{x,y}$), and range of neighbor nodes ($R_{nn}$). Only the sensors whose distance value is less than or equal to range of neighboring node is considered as the neighboring nodes (Line-7). The next part of the algorithm performs computation of the vulnerability metric $vul$ considering sensors ($n$) and position of node ($p_{x,y}$) (Line-8). It does so by constructing a function $f_2(x)$ that carry out estimating the Euclidean distance between the normal node and the adversary node followed by summation of all the estimated distance. The next part of the algorithm implementation performs computation of vicinityfactor (Line-9). The algorithm uses a function $f_3(x)$ that computes the distance between the regular and malicious node and checks if they are located by closely to each other.

2.4. Applying rule-set to resist adversary

The algorithm then assesses the primary optimization by applying the primary rule-set $\Phi_P$ on initialized energy and number of neighbor nodes. According to this primary optimization, a regular node should spontaneously dissipate energy. However, a malicious node will always try to save maximum energy until and unless it has not initiated its attack. It will mean that if a malicious node has not launched an attack, it is not feasible to differentiate attacker node and normal node. Hence, the presence of abnormally higher
amount of energy will be one of the indications of presence of malicious node in primary optimization process. Establishment of communication with such nodes is instantly aborted after this. However, the algorithm extends the primary optimization to the secondary optimization in order to further ascertain about the efficiency of the resistance. In the secondary optimization, the algorithm considers applying secondary rule-set $\phi$, to finally ensure that no such vulnerable node is selected in this process. The function takes the input arguments to ensure that only the secured clusterhead is selected in this process (Line-10). Finally, the algorithm selects the highly utilized secondary optimized value that ends up selecting the most eligible clusterhead $n_{c3}$ (Line-11).

2.5. Logic of identification and resistance

The proposed system applies a unique logic of identifying and conforming the presence of malicious node and implements a mechanism to resist any form of intrusion. The identification process of the adversary is carried out in dual steps. In the first step of identification, the sensor performs exchange of information in the form of beacons with each other that also bears a specific field for residual energy. The proposed system assumes that all the clusterheads are higher degree of communication synchronocity with each other by periodic exchange of beacons among the clusterheads during inter-cluster communication. According to this scheme, a cluster head bears all the authenticated information related to its associated member nodes and this will call for eventually monitoring all the resource-related information too. Even if a conventional TDMA scheme is assumed to be implemented than the time for data fusion as well as aggregation is almost fixed. This entails that regular nodes spontaneously deplete energy in the progressive step of data aggregation and hence in this process, if there is a presence of any node that bears abnormally more residual energy than it is less likely that such node will be regular node. However, presence of such node with higher residual energy cannot be used to conclude that the node is a malicious one and therefore, it proceeds for next round of check. For the purpose of confirming the presence of malicious nodes at the end of preliminary checking stage, the algorithm considers multiple distance-based parameters e.g. vulnerability and vicinity parameters in order to carry out more indepth investigation about the behaviour of the sensors. The underlying principle of this resistance theory is that a malicious node will not launch any form of attack when it joins a new network. At the same time, it will neither reside in a same network for a longer period of time. The gain of the malicious node is only at the instance when they are successful in their attempt. For this purpose, the adversary will be required to increase its attacking event as far as feasible. The next step is to check for the presence of intermittent communication links, which directly depicts that this link leads to malicious nodes at some point.

Therefore, as per the logic of the proposed algorithm, the presence of more number of consistent remnant energy is a direct representation of the fact that there is a malicious node at the end. However, there is also a possibility that it could be regular node. In such case, the prior history from the routing table will be checked. If upon checking it is found that that node has transmitted many data in past and still it retains more energy than it is a positive indication of malicious node or else it could be regular node. At the same time, if the number of distances that are always increasing in its size than it directly depict that there is a spontaneous spread of attack which is feasible by capturing the public keys. Therefore, it is imperative that normal sensors will always have less number of increasing spatial distances (as it will drain more energy) as well as increasing distance among different sensors. On the other hand, it is less likely that adversary will stop after compromising few nodes as fair chances are there to increase the number of attacks. Once the malicious node is confirmed for its presence than all the communication to and from the adversary will be eliminated and subsequent a flag message about the node and its respective neighbors are updated to all the nodes as an update. Hence, the proposed algorithm offers significantly cost-effective solution to offer significant resistance to any lethal forms of attack in WSN. The next section discusses about the outcomes obtained after implementing this algorithm.

3. RESULT ANALYSIS

This section discusses about the outcomes obtained after implementing the algorithm discussed in prior section. The complete emphasis of the analysis is given to the mainly two performance scale i.e. throughput and energy. Scripted in MATLAB, the analysis of the proposed algorithm is carried out considering 500-1000 sensors considering large scale network of area 1000x1200 m². The sensors are programatically designed bearing the standard characteristic of MEMSIC nodes along with 2500 bits of packet and 0.5 Joules of energy. The study outcome of the proposed system is compared with the standard security protocol of SecLEACH [35] because this protocol is known for both its security benefits and energy conservation in WSN.
The outcome exhibits in Figures 4 and 5 highlights that proposed system offers good energy conservation in WSN. A closer look into Figure 4 highlights that proposed system offers good network lifetime as there are good number of sustaining nodes over increasing number of simulation rounds as compared to conventional SecLEACH algorithm. The prime reason behind this is SecLEACH uses a typical encryption mechanism that uses recursive function in order to generate the secret key for authentication. This phenomenon will require a node to always have specific amount of resources to carry out this task. Moreover, the placement of the base station is another bigger challenge in existing system that result in massive traffic converging using single hop mainly. Hence, SecLEACH cannot offer energy conservation as the test environment chosen is a combination of active attacks with dynamic sensor nodes that cannot be supported energy efficiently by SecLEACH.

Figure 6 discusses the comparative analysis of the throughput for both proposed and SecLEACH algorithm. The outcome shows that SecLEACH couldn’t offer sufficient throughput more than 40% completion of simulation rounds, whereas proposed system offers spontaneously increased amount of throughput till the 90% of the simulation rounds proving that proposed system offers significant throughput performance along with optimal security. Figure 7 highlights that proposed system offers significantly less energy fluctuation as compared to SecLEACH. Increased fluctuation in energy is always linked with lack of synchronicity among the nodes that results in failure of an effective data transmission as well as non-supportability of any form of energy-efficient algorithms. Increased energy variance of SecLEACH is caused due to lack of synchronous as well as lack of supportability of multi-hop communication scheme. This results in more degradation of energy but in a highly intermittent way. On the other hand, the proposed system offers increasing events of dissemination of the updates that results in exchange of nearly updated data everytime thereby assisting the other sensors know about the event of successful data transmission or presence of malicious node. Hence, the proposed system is found to offer more resistivity against maximum form of threats with lesser degradation to the energy consumption among the sensor nodes.

<table>
<thead>
<tr>
<th>Figure 4</th>
<th>Figure 5</th>
<th>Figure 6</th>
<th>Figure 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>[Image]</td>
<td>[Image]</td>
<td>[Image]</td>
<td>[Image]</td>
</tr>
</tbody>
</table>

Figure 4. Comparative analysis of alive nodes
Figure 5. Comparative analysis of residual energy
Figure 6. Comparative analysis of throughput
Figure 7. Comparative analysis of variance
4. CONCLUSION

This paper has discussed that identifying a malicious node is not at all an easy task that can also justify the reason behind not 100% fail—proofness about the existing security approaches in WSN. The logic of the proposed method claims that there is always a common and a typical behavior of any adversary node, which saves the time to investigate multiple attacks. In this regard, the energy acts as an essential resource for all the sensors whether it is regular node or malicious node. The proposed system considers dual level of filtering the patterns where in the first level an efficient clusterhead is selected while in the second level security characteristics is emphasized more. A common track of retention of abnormal residual energy is considered as one of the criteria of malicious node whereas there are also dual checks to confirm this fact. The simulation outcome shows that it is practically feasible to offer good level of security to majority of attacks as it offers good network lifetime and data delivery performance in comparison to existing secure and energy-efficient approaches.

REFERENCES


BIOGRAPHIES OF AUTHORS

Tejashwini N., she has completed her B. E from SVCE, Bengaluru Karnataka and M. Tech is from Dr. AIT, Bengaluru, Karnataka, she having six years of teaching and research experience in respective engineering colleges under Visvesvaraya Technological University, Belagavi, Karnataka, India. Her interest includes Digital image communication, Wireless communication, Wireless sensor network and Operating system. Currently she is pursuing her research for PhD under Visvesvaraya Technological University, Belagavi, Karnataka. She has published her work in various international conferences and journals.

D. R. Shashi Kumar, Currently working as a Professor and HOD, Department of CSE, Cambridge Institute of Technology, Bengaluru, Karnataka, India. He has more than 29 working in various capacities. He has More than 20 Research Papers (National and International) in his credit and has chaired national and international conferences. He has published 10 research papers in refereed International Technical Journals with good Impact factor. His research interest area is Digital Image Processing, Computer Networks- Specialization Microprocessors, Data Mining, and Neural Networks.

K. Satyanarayana Reddy, Currently working as a Professor and HOD, Department of ISE, Cambridge Institute of Technology, Bengaluru, Karnataka, India. He has more than 25 years of experience in academics in the field of Computer Science. He has more than 25 Research Papers (National and International) in his credit and has chaired national and international conferences. Delivered Keynote address in few national level conferences. His area of interest is Artificial Intelligence and Robotics, Computer Networks, Data Communications, Software Engineering, Theoretical Computer Science, Wireless Sensor Networks.