Abstract—In the presence of collision, secure data aggregation technique for wireless sensor networks, due to limited computational power and energy resources of sensor nodes, aggregation of data from multiple sensor nodes done at the aggregating node is usually accomplished by simple methods such as averaging. In several sensor applications, the data collection will form from the entity nodes be aggregate next to a base station or host computer. The performance of the network aggregation of data sensor at intermediary nodes enrooted towards the base station. For the duration of the process of the collection of the data in network sensor for energy outgoings of the nodes sensor and the communication in the clouds are reduces for the hop-by-hop data aggregation, this aggregation are incredibly vital technique. Iterative filtering algorithms enhanced grand undertake intended for such an intention. In this paper we assume that the aggregator itself is not compromised and concentrate on algorithms which make aggregation secure when the individual sensor nodes might be compromised and might be sending false data to the aggregator. We assume that each data aggregator has enough computational power to run an IF algorithm for data aggregation. First we apply Iterative Filtering algorithm to the Aggregator nodes also to provide more security and assign unique values to all the nodes using SHA algorithm.

Keywords—Data Aggregation, Collusion Attacks, Iterative Filtering Algorithm, SNA Algorithm, Wireless Sensor Networks.

I. INTRODUCTION

Security breach can ensue in a sensor network not only while relaying in order to the end-user but also though generating information. The ability of a sensor network to perform its task depends not only on its capability to communicate between the nodes, although also on its capability to intellect the physical environment and collectively progression the sense data. This decentralized in network decision-making, which relies on the inherent trust among the nodes, can be abused by adversaries to carry out security breaches through compromised nodes. Reminder that sensor nodes are envisioned to be low-cost which make it infeasible for manufacture to construct them tamper-resistant; an opponent be able to undetectably take the control of a sensor node by physically compromising it. An opponent can then potentially introduce out of order data or decisions to take in the complete network. Cryptographic and authentication mechanism alone cannot be use to resolve that problem as internal adversarial nodes will have entrance to suitable cryptographic keys. In addition spiteful attacks, sensor nodes are also susceptible to organization fault. Non-malicious behaviour such as not working of radio/sensors can also result in the generation of bogus data, bringing equally detrimental effects to the performance of the network. The extremely nature of this type of misbehaviour is external the empire of cryptography [10].

Wireless sensor networks are usually composed of hundreds or thousands of inexpensive, low-powered sensing devices with limited memory, computational, and communication resources. These networks offer potentially low-cost solutions to an array of problems in both military and civilian applications, including battlefield surveillance, target tracking, environmental and health care monitoring, wildfire detection, and traffic regulation. Due to the low deployment cost
requirement of wireless sensor networks, sensor nodes have simple hardware and severe resource constraints. Hence, it is a challenging task to provide efficient solutions to data gathering problem [5]. Data aggregation protocols aim to combine and summarize data packets of several sensor nodes so that amount of data transmission is reduced. An example data aggregation scheme is presented in Figure 1 where a group of sensor nodes collect information from a target region. When the base station queries the network, instead of sending each sensor node’s data to base station, one of the sensor nodes, called data aggregator, collects the information from its neighbouring nodes, aggregates them (e.g., computes the average), and sends the aggregated data to the base station over a multi hop path. As illustrated by the example, data aggregation reduces the number of data transmissions thereby improving the bandwidth and energy utilization in the network.

![Figure 1: WSN structure](image)

Trust and reputation systems (abbreviated TRS hereafter) represent an important class of decision support tools that can help reduce risk when engaging in transactions and interactions on the Internet. From the individual relying party’s viewpoint, a TRS can help reduce the risk associated with any particular interaction. From the service provider’s viewpoint, it represents a marketing tool. From the community viewpoint, it represents a mechanism for social moderation and control, as well as a method to improve the quality of online markets and communities. The same fundamental principles for creation and propagation of trust and reputation in traditional communities are too used by online TRSs. The major variation is that trust and reputation formation in traditional communities typically is comparatively inefficient and relies on physical communiqué example that through word-of-mouth, while online TRSs are support by particularly well-organized networks and computer systems. In theory it is possible to intend very effective trust and reputation management in online communities, but the consistency of computed trust and reputation scores, and thereby the usefulness of the TRS itself, moreover depends going on the robustness of the TRS [6]. The reputation mechanism, however, opens up new vulnerabilities as the raters may present unreliable or malicious reports, demonizing the reputations of the service providers incorrectly [5].

1.1. Motivation
The data aggregator itself is not compromised and concentrate on algorithms which make aggregation secure when the individual sensor nodes might be compromised and might be sending false data to the aggregator and assume that each data aggregator has enough computational power to run an IF algorithm for data aggregation. Nodes are also to provide more security.

1.2. Contribution
We enhanced a secure data aggregation technique for wireless network in the presence of the collusion attack by proposed algorithms and based on some parameter matrices such as delay, RMS error and Robust aggregate Affine.
1.3. Organisation
The rest of the paper is organised as follows. A brief literature of related work is presented in section 2. Background of the proposed work as shown in section 3. Section 4 defines the problem and objectives. The methodology and proposed algorithm as discussed in section 5. Section 6 represents the experimental results and performance parameter. And section 7 concludes the proposed work.

II. LITERATURE SURVEY

Yan Sun et al [1] in this method, they have designed a trust-based framework for data aggregation with fault tolerance by extends Josang’s trust form. Proposed scheme can significantly decrease the impact of erroneous data and provide assessable trustworthiness for aggregated data. General simulations and experiment on a real tested are conduct to confirm the efficiency of the planned framework on both continuous media stream and separate data. Hence, it is practical for multitask in wireless multimedia sensor networks. In the potential, they analyze more data interruption form and design more perfect fault tolerant and intrusion-tolerant system.

E. Ayday et al[2] this method proposed “An iterative algorithm for trust and reputation management,” a small different iterative algorithm. Their main differences from the other algorithms. Ratings have a time- is count factor, so in time, their importance will fade out; and the algorithm maintain a blacklist of users who are particularly bad raters. Though the existing IF algorithms regard as easy immoral behavior by adversaries, none of them takes into account sophisticated malicious scenarios such as collusion attacks.

L.-A. Tang et al [3] this method proposed a trust framework for sensor networks in cyber physical systems such as a battle-network in which the sensor nodes are working to detect imminent enemies and send alarms to a command center. Although fault detection problems have been addressed in apply the trust and reputation systems in the above research, nobody of them take into account complicated collusion attacks scenarios in adversarial environments. Reputation and trust concepts can be used to overcome the compromise the node detection and secure data aggregation problems in WSNs.

S. Ozdemir et al [4] this method Proposed “Integration of false data detection with data aggregation and confidential transmission in wireless sensor networks.” The major scheme of forged aggregator detection in the scheme proposed in is to employ a number of monitoring nodes which are running aggregation operations and providing a MAC value of their aggregation results as a part of MAC in the value computed by the cluster aggregator. High computation and transmission cost necessary for MAC-based integrity checking in this scheme makes it unsuitable for deployment in WSN.

Suat Ozdemir et al I[5] They explained an association between security and data aggregation method in wireless sensor networks. Taxonomy of secure data aggregation protocols is given by survey the present “state-of-the-art” work in this area. Based on the existing research, the open research areas and future research directions in secure data aggregation concept are provided. They have provides a complete re-evaluate of secure data aggregation model in wireless sensor networks.

Audun Josang et al [6] They have propose the trust and reputation systems is make stronger then the value of markets and communities by providing an incentive for good behaviour and value of services, and through sanction bad performance with low value services. presently, robustness analysis of TRSs is mostly done throughout easy replicated scenario implement with the TRS designers themselves, and that cannot be consider the dependable confirmation for how those systems would carry out in a sensible situation. In order to put robustness needs it is significant to be on familiar terms with how vital robustness really is in an exacting community or market. They have discusses research challenges for trust and reputation systems, and propose a investigate schedule used for developing sound and reliable robustness principles and mechanisms for trust and reputation systems.
Kevin Hoffman et al [7] this method, proposed a Survey of attacks on Reputation Systems. Firstly they have focusing on the design dimensions of reputation systems and the equivalent attack and defences and developed an analysis framework this preserved to be used when common criteria for evaluating and compare reputation systems. They defined an attacker model with classify recognized and other potential attack on reputation systems within this model. Defence mechanism and their equivalent strength and weakness were discussed. And they demonstrated the rate of the investigation framework and attack and defence characterization with survey some key reputation systems, drawing insight base on the novel framework. This test framework is also valuable for future research in that it provides understanding into the implications of design choices.

H. Chan et al[8] proposed “Secure hierarchical in-network aggregation in sensor networks,” here they have studies focus on detecting false aggregation operations by an adversary, that is, on data aggregator nodes obtaining data from source nodes and producing wrong aggregated values. Consequently, they address neither the problem of false data being provided by the data sources nor the problem of collusion. However, when an adversary injects false data by a collusion attack scenario, it can affects the results of the honest aggregators and thus the base station will receive skewed aggregate value.

Yi Yang et al [9] they have propose SDAP, a Secure Data Aggregation Protocol for large-scale sensor networks. With using divide and conquer, and partition the aggregation tree into groups to decrease the significance of high-level nodes in the aggregation tree. They use commit and attest so that the base station has a way to confirm the aggregates. In further enrich the protocol in additional feature.

S. Ganeriwal et al [10] in this method, proposed a general reputation framework for sensor networks in which each node develops reputation estimation for new nodes through observe its neighbors which compose a trust community for sensor nodes in the network. Proposed a trust based framework which employs correlation to detect damaged readings. additionally, they have introduce a ranking structure to associate a level of trustworthiness by means of each sensor node based on the number of neighboring sensor nodes are sustaining the sensor.

III. BACKGROUND

For the sensor network topology, the sensor nodes are divided into disjoint clusters, and each cluster has a cluster head which acts as an aggregator. Data are periodically collected and aggregated by the aggregator. In this paper we assume that the aggregator itself is not compromised and concentrate on algorithms which make aggregation secure when the individual sensor nodes might be compromised and might be sending false data to the aggregator. We assume that each data aggregator has enough computational power to run an IF algorithm for data aggregation and SNA algorithm. An IF algorithm for computing reputation of objects and raters in a rating system. We briefly describe the algorithm in the context of data aggregation in WSN and explain the vulnerability of the algorithm for a possible collusion attack. We note that our improvement is applicable to other IF algorithms as well.

III. PROBLEM DEFINITION AND OBJECTIVES

Proposes a solution for such vulnerability by providing an initial trust estimate which is based on a robust estimation of errors of individual sensors. When the nature of errors is stochastic, such errors essentially represent an approximation of the error parameters of sensor nodes in WSN such as bias and variance. However, such estimates also prove to be robust in cases when the error is not stochastic but due to coordinated malicious activities. Such initial estimation makes IF algorithms robust against described sophisticated collusion attack, and, we believe, also more robust under significantly more general circumstances; for example, it is also effective in the presence of a complete failure of some of the sensor nodes.

We assume that the aggregator itself is not compromised and concentrate on algorithms which make aggregation secure when the individual sensor nodes might be compromised and might
be sending false data to the aggregator. We assume that each data aggregator has enough computational power to run an IF algorithm for data aggregation. Nodes also to provide more security.

- Apply Iterative Filtering algorithm to the Aggregator nodes also to provide more security.
- Assign unique values to all the nodes using SHA algorithm.

IV. EXISTING SYSTEM

Iterative Filtering (IF) algorithms are an attractive option for WSNs because they solve both problems data aggregation and data trustworthiness assessment using a single iterative procedure. Such trustworthiness estimate of each sensor is based on the distance of the readings of such a sensor from the estimate of the correct values, obtained in the previous round of iteration by some form of aggregation of the readings of all sensors. Such aggregation is usually a weighted average; sensors whose readings significantly differ from such estimate are assigned less trustworthiness and consequently in the aggregation process in the present round of iteration their readings are given a lower weight.

It is particularly dangerous once launched against WSNs for two reasons. First, trust and reputation systems play critical role in WSNs as a method of resolving a number of important problems, such as secure routing, fault tolerance, false data detection, compromised node detection, secure data aggregation, cluster head election, outlier detection Second, sensors which are deployed in hostile and unattended environments are highly susceptible to node compromising attack. While offering better protection than the simple averaging, our simulation results demonstrate that indeed current IF algorithms are vulnerable to such new attack strategy.

Disadvantages

a) Trust and reputation systems play critical role in WSNs as a method of resolving a number of important problems, such as secure routing, fault tolerance, false data detection, compromised node detection, secure data aggregation, cluster head election, outlier detection.

b) Sensors which are deployed in hostile and unattended environments are highly susceptible to node compromising attack.

V. METHODOLOGY

Figure 2 represents the block diagram of proposed method; it consists of sensor nodes, cluster formation, robust data aggregation and sink. Whereas a sensor nodes block has an ‘n’ number of nodes and clusters are formed with the ‘n’ number of clusters then it subjected to the robust data aggregation. The data aggregation is a technique used to solve the implosion and overlap problems in data centric routing. Data coming from multiple sensor nodes is aggregated as if they are about the same attribute of the phenomenon when they reach the same routing node on the way back to the sink. Data aggregation is a widely used technique in wireless sensor networks. The security issues like data confidentiality and integrity, in data aggregation become vital when the sensor network is deployed in a hostile environment. In wireless sensor networks (WSNs), all the data collected by the sensor nodes are forwarded to a sink node. Therefore, the placement of the sink node has a great impact on the energy consumption and lifetime of WSNs.
5.1. Sensor node
A sensor node, also known as a mote, is a node in a sensor network this is capable of performing a few processing, gathering sensory information and communicating with other connected nodes in the network. A mote is a node but a node is not always a mote. The sensor networks are intense wireless networks of small, low-cost sensors, which disseminate and collect data from environmental. WSNs are facilitating to monitor and scheming of physical environments from remote locations with better accuracy rate. We have many applications for WSN in a variety of fields such as environmental monitoring, military purposes and gathering sensing information in inhospitable locations. Sensor nodes have a variety of energy and computational constraints because of their inexpensive nature and Adhoc method of consumption. A sensor network is an infrastructure comprised of sensing (measuring), computing, and communication elements this gives an administrator to the ability to instrument, view, and react to events and phenomena within a specific surroundings. Typical applications include, but are not limited to, data collection, monitoring, surveillance, and medical telemetry. During, adding up to sense, individual is often also involved in manage and establishment. There are few basic components in a sensor network. An assembly of distributed or localized sensors, an interconnecting network (generally, but not forever, wireless-based), a central point of information clustering and a set of computing resources at the central point to handle data relationship, event trending, status query, and data mining. Thus, the sensing and computation nodes are considering component of the sensor network. In information, some of the compute may be done in the network itself. Because of the potentially large amount of data will collected, algorithmic method for data organization take part in an main role in sensor networks.

5.2. Cluster formation
The cluster will form a ‘n’ numbers of the clusters as shown in figure2. Randomly deployed sensor networks require a cluster formation protocol to separation the network into clusters. When cluster heads are required, nodes in each cluster may also perform a organizer determination protocol to establish their cluster. Several cluster formation have been planned for wireless sensor networks. In Cluster-First approaches the sensor nodes first form clusters and each cluster then elect its cluster. Those approaches need every node in one cluster consent on the same relationship before electing their cluster, and sensor nodes are almost always separated into clique so those nodes in each clique be able to in a straight line communicate with one after other. Mainly, existing cluster formations assume kind environment, and can’t live attack from malevolent participant in aggressive environments. In the Leader approaches, malicious nodes may lie regarding
their metrics (e.g., increase transmission power for cluster-head advertisement messages) to make themselves elected as cluster. Since a consequence, they know how to manage all the nodes in their clusters. correspondingly, not any of the Cluster-First protocols can assurance a reliable view on clique memberships and two secure leader selection algorithms by using a trusted authority to certify each node’s metrics used in the leader election process. Though, those algorithms are suppose to do all participate nodes are reliable and no messages are lost or delayed, which cannot be guaranteed when there are malicious nodes.

5.3. Robust data aggregation
The data aggregation is mainly; to increase the network life span by reduces the resource consumption of sensor nodes. While increasing network lifetime, data aggregation protocols may degrade significant value of service metrics in wireless sensor networks, such as data accuracy, latency, fault-tolerance, and security. Hence, the design of an efficient data aggregation protocol is a naturally demanding task because the protocol designer must trade off between energy efficiency, data accuracy, latency, fault-tolerance, and security. Here command to achieve this trade off, data aggregation techniques are strongly together with how packets are routed throughout the system. Thus, the construction of the sensor network plays a vital role in the performance of different data aggregation protocols.

5.3.1. Iterative filtering
In organize to adjust the operation of IF algorithms adjacent to the aforesaid attack state, our proposed approach provide a robust initial inference of the trustworthiness of sensor nodes to be used in the first iteration of the IF algorithm. Most of the conservative statistical evaluation methods for variances occupy by use of the sample mean. For this reason, propose a robust variation inference technique in the case of slanted example mean is vital element of our methodology.
(a) Identification of a new complicated and powerful attack adjacent to IF based reputation systems which reveal a rigorous susceptibility in iterative filtering algorithms.
(b) A new method for fault estimation of sensors nodes which is effectual in a wide range of sensor fault and not vulnerable to describe the attack.
(c) Design of an efficient and robust aggregation technique encouraged by the Maximum Likelihood Estimation (MLE), which utilize an estimation of the noise parameters obtained with contribution 2 beyond.
(d) Enhanced IF schemes able to protect adjacent to complicated collusion attacks by given that an initial estimate of trustworthiness of sensors using inputs from contributions 2 and 3 above;
(e) A novel collusion detection and revocation method based on an initial approximation of the aggregate values as well as distribution of differences of each sensor reading values. The below algorithm describes an iterative filtering.

Algorithm 1: Iterative Filtering Algorithm

Input: X, n, m.
Output: The reputation vector r
l ← 0;
w\(^{(0)}\) ← 1;
repeat
Compute r\(^{(l+1)}\);
Compute d;
Compute w\(^{(l+1)}\);
l ← l + 1;
Until reputation has converged;

Algorithm 2: SHA Algorithm

Step1: Append padding bits
Message is padded with a 1 and as many 0’s as necessary to bring the message length to 64 bits less than an even multiple of 512.

**Step2: Append length**

64 bits are appended to the end of the padded message. These bits hold the binary format of the 64 bits indicating the length of the original message.

### 5.3.2. Maximum likelihood estimation

The method of maximum likelihood corresponds to many well-known estimation methods in statistics. For example, one may be interested in the heights of adult female penguins, but be unable to measure the height of every single penguin in a population due to cost or time constraints. Assuming that the height are usually some unknown mean and variance, the mean and variance can be predictable with MLE while only knowing the height of some example in general population. MLE would complete this by attractive the mean and variance as parameters and find the particular parametric values that build the practical results the mainly likely known the model.

MLE is not as widely familiar along with modelers in psychology, but it is a standard approach to parameter estimation and inference in statistics. MLE has many optimal properties in estimation: sufficiency (complete information about the parameter of interest contained in its MLE estimator); consistency (true parameter value that generate the data improved asymptotically, i.e. for data of sufficiently big samples). The method of maximum likelihood provides estimators that contain equally a reasonable intuitive basis and many desirable statistical properties. The technique is very broadly applicable and is easy to relate. Once a maximum-likelihood estimator is resulting, the common theories of maximum-likelihood estimation provide typical error, arithmetical tests, and other results useful for statistical conclusion.

### 5.3.3. Noise parameter estimation
Noise is present in all images captured by image sensors. Due to photon emission and photoelectric effects that are the foundations of the ways in which quantum mechanics enable image sensors, in fact, random noise is a essential wicked of image sensors so as to continue their require our attention. The goal of this work is to provide a wide-ranging classification of random noise in behavior that improves post-image-capture signal giving out steps. We obtain the Poisson approximation to model the measurement noise that is the result of photon arrival and photon recapture. A novel methodology to study the parameters that explain the noise is residential. We terminate by present preliminary verification that exact noise model would get better image de-noising, particularly in the low photon count/high noise regime.

Figure 3 represents the overall control flow of proposed methodology; start with the initialize the sensor nodes with some error in the network. These sensors can be modeled by using a Gaussian variable to estimate bias. Whereas the estimate bias value are use to defining and computing matrices to estimate variance. By using maximum likelihood estimation, reputation vector is estimated, then reputation the number of iteration and forwards to the sink (base station) and also reduces the number of iterations, these iteration is passes to the sink.
VI. EXPERIMENTAL RESULTS AND PARAMETERS

Figure 4: Sensor nodes with some error

Figure 5: cluster formation

Figure 6: cluster 1

Figure 7: cluster 2

Figure 8: cluster 3

Figure 9: cluster 4

Figure 10: cluster 5

Figure 11: aggregate nodes

Figure 12: selection of base station node

Figure 13: identifies the misbehavior nodes

Figure 14: node 7(cluster head) is misbehavior

Figure 15: 11, 19, 27, 35 nodes are misbehavior

Figure 4 represents the initializing the nodes to sending the messages to each node. To finding the misbehavior nodes and trusted nodes, because it will assume the data’s are secured, when it will communicating the messages from one to another sensor nodes. Figure 6, 7, 8, 9, 10 gives a formation of five clusters, then we select a aggregation nodes represents in Figure 11 and identifies the misbehavior nodes as shown in figure 13, 14, 15. Figure 16 represents a plot graph of delay, RMS, Robust aggregation affine.

4.1. Performance parameters:

a) Delay: The delay of a network specifies how much time it takes for a bit of data to travel across the network from one node or base station to another node.

b) RMS error: The Root Mean Square Error is a frequently used measure of the difference between values predicted by a model and the values actually observed from the environment that is being modeled. These individual differences are also called residuals, and the RMSE serves to aggregate them into a single measure of predictive power.
The RMSE of a model prediction with respect to the estimated variable $X_{model}$ is defined as the square root of the mean squared error

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$

Where $X_{obs}$ is observed values and $X_{model}$ is modeled values at time/place $i$.

c) **Robust aggregate Affine**: the robust aggregate affine graph as shown in figure 16 (c), gives a data secure of the aggregation of proposed method.
VII. CONCLUSION

In this paper introduced a secure data aggregation technique for wireless sensor networks in the presence of collusion attacks. Moreover, we proposed an enhancement for the IF algorithms with providing an initial estimate of the trustworthiness of sensor nodes which makes the algorithms not just collusion robust, but as well more accurate and faster converging. Based on the survey the relationship between data aggregation and security requirement are explained. We have presented on wireless sensor networks for the secure data aggregation. Here we improves the data aggregation on IF algorithm and SHA algorithm for the trustworthiness of the sensor node, where SHA algorithms are the used to take a all nodes to be unique values and Apply Iterative Filtering algorithm to the Aggregator nodes also to provide more security, for security basis. And this proposed method is using NS-2 simulation and performs a parameters of delay, RMS-error and Robust aggregate affine. In future we can investigate to protect against compromised aggregators and also can implement the deployed sensor network. And also further improve the protocol in details like breadth based attestation and content based attestation technique to the used protocol.

REFERENCES

2. Erman Ayday, “Iterative Algorithms for Trust And Reputation Management And Recommender Systems”. 
3. Lu-An Tang , Xiao Yu, Sangkyum Kim,Quanquan Gu, Jiawei Han, Alice Leung and Thomas La Porta, “Trustworthiness Analysis of Sensor Data in Cyber-Physical Systems”.
13. Hao Liao, Giulio Cimini Mat_u_s Medo,” Measuring quality, reputation and trust in online communities”.

Figure 16: (a) Delay (b) RMS-error (c) Robust Aggregate-Affine