Energy Aware Load-Balancing Of Parallel Mining Of Frequent Sequences.

B.Manohari, Dr Mala,

Research Scholar, Department of Computer Science, Marudupandiyar College, Thanjavur
Department of Computer Science, Marudupandiyar College, Thanjavur

Abstract

Due to a need for toughness of observe and low cost of the nodes, wireless sensor networks (WSNs) are usually unneeded. Data from multiple sensors is combined at an aggregator node which then ahead to the base stations only the combined values. At present, due to limits of the computing power and power resource of sensor nodes, data is aggregated by tremendously simple algorithms such as averaging. However, such aggregation is known to be very exposed to faults, and more importantly, spiteful attacks. This cannot be preparation by cryptographic methods, because the attackers generally gain complete access to information stored in the negotiation nodes. For that reason data aggregation at the aggregator node has to be accompanying by an evaluation of trustworthiness of data from individual sensor nodes. Thus, better, more complicated algorithms are needed for data aggregation in the future WSN.

I. INTRODUCTION

A trustworthiness evaluation at any given moment represents a combined of the behavior of the participants up to that instant and has to be vigorous in the presence of various types of faults and malicious behavior. There are a number of enticements for attackers to influence the trust and status scores of participants in a circulated system, and such manipulation can strictly impair the performance of such a system. The main goal of malicious attackers is aggregation algorithms of trust and status systems. Trust and status have been recently suggested as an effective security mechanism for Wireless Sensor Networks (WSNs).

Although sensor network are being more and more deployed in many application domains, assessing trustworthiness of statement data from distributed sensors has stay behind a challenging issue. Sensors deployed in aggressive environments may be subject to node cooperation attacks by opponent who mean to bring in false data into the system. In this context, assess the trustworthiness of the together data and announcing decision makers for the data trustworthiness turn out to be a challenging task. As the computational authority of very low power processors radically increases, mostly driven by demands of mobile computing, and as the cost of such technology drops, WSNs will be able to have the funds for hardware which can implement more complicated data aggregation and trust assessment algorithms; an example is the recent appearance of multi-core and multi-processor systems in sensor nodes. Iterative Filtering (IF) algorithms are an eye-catching option for WSNs because they solve both problems - data aggregation and data trustworthiness evaluation - using a single iterative method. Such trustworthiness estimate of each sensor is based on the distance of the readings of such a sensor from the approximation of the acceptable values, attain in the previous round of iteration by some form of aggregation of the interpretation of all sensors. Such aggregation is typically a weighted average; sensors whose readings considerably differ from such estimate are assigned less trustworthiness and as a result in the aggregation process in the present surrounding of iteration their readings are given a lower weight. In recent years, there has been an increasing amount of writing on IF algorithms for trust and status systems.
2. Wireless Sensor Networks (WSNs)

First, trust and status systems play serious role in WSNs as a method of resolving a number of significant problems, such as protected routing, responsibility tolerance, fake data detection, negotiation node detection, secure data aggregation, gather head election, outlier exposure, etc. Second, sensors which are organize in aggressive and unattended situation are highly vulnerable to node conciliation attacks. While offering better shield than the simple averaging, the replication results display that certainly current IF algorithms are susceptible to such new attack approach As it will see, such vulnerability to complicated involvement attacks comes from the fact that these IF algorithms begin the iteration process by giving an equal trust value to all sensor nodes. In this work, it suggests a resolution for such vulnerability by providing an initial trust approximation which is based on a robust opinion of errors of individual sensors. When the nature of faults is stochastic, such faults essentially symbolize an approximation of the error constraint of sensor nodes in WSN such as bias and dissent. However, such estimates also establish to be robust in cases when the error is not stochastic but due to synchronized malicious activities.

Identification of a new difficult collusion attack against IF based status systems which exposes a severe vulnerability of IF algorithms;

- A novel method for opinion of sensors’ errors which is efficient in a wide range of sensor errors and not disposed to the illustrate attack;

- Design of an resourceful and robust aggregation method encouraged by the MLE, which make use of an estimate of the noise parameters.

- Enhanced IF schemes (LBS) Level Based Scheme competent to protect against complicated collusion attacks by providing an initial estimate of trustworthiness of sensors using inputs from assistance 2 and 3 above;

- It provides a thorough experimental evaluation of effectiveness and competence of the proposed aggregation method. The results explain that the technique provides both higher accurateness and better collusion confrontation than the existing manners.

3. REVIEW OF LITERATURE

Ankit Tripathi, Sanjeev Gupta, Bharti hourasiya, 2014, presented a paper An Energy-Aware Spanning Tree Algorithm for Data Aggregation in Wireless Sensor Networks. In wireless sensor networks, data aggregation is used to collect local data from adjacent nodes (sources) and produce a data report. In this article, they propose E-Span, which is an energy-aware spanning tree algorithm. E-Span is a distributed protocol and facilitates the sources within an event region to execute data aggregation. In E-span, the source node which has the highest outstanding energy is chosen as the root. Other source nodes choose their equivalent parent node among their neighbors based on the remaining energy and distance to the root.

Ming-Jer Tsai, 2013, presented a paper was Survey on Data Aggregation Techniques for Wireless Sensor Networks. Wireless sensor networks (WSNs) consist of sensor nodes. These networks have enormous function in territory monitoring, disaster management, security and military, etc. Wireless sensor nodes are very small in size and have limited processing capability and very low battery power. This limit of low battery power makes the sensor network level to failure. Data aggregation is a very key technique in WSNs. Data aggregation helps in reducing the energy consumption by eliminating idleness. This work focuses on summarizing different approaches used for the purpose of data aggregation and its various energy-efficient uses in WSN.
Deepali Virmani, Tanu Sharma & Ritu Sharma, 2012 has introduced Adaptive Energy Aware Data Aggregation Tree for Wireless Sensor Networks. To gather the difficulty of wireless sensor networks (WSNs) where data are typically aggregated at a single resource prior to transmitting to any distant user, there is a need to establish a tree structure inside to aggregate data. In this article, an adaptive energy aware data aggregation tree (AEDT) is proposed. The proposed tree uses the maximum energy available node as the data aggregator node. The tree incorporates sleep and awake technology where the communicating node and the parent node are only in awake state rest all the nodes go to sleep state saving the network energy and attractive the network lifetime.

Sanjeev SETIA, Sankardas Roy J, 2012 has proposed a new system in Secure Data Aggregation in Wireless Sensor Networks. In a lot of sensor applications, the data collected from individual nodes is aggregated at a base station or host computer. To reduce energy consumption, many systems also perform in-network aggregation of sensor data at middle nodes on the road to the base station. Most existing aggregation algorithms and systems do not include any requirements for security, and as a result these systems are in danger to a broad range of attacks. In particular, compromised nodes can be used to add false data that leads to incorrect aggregates being computed at the base station.

Ali Norouzi1, Faezeh Sadat, 2012, presented a study on Group-Independent Spanning Tree for Data Aggregation in Dense Sensor Networks. Today, there survive many algorithms and protocols for constructing aggregation or distribution trees for wireless sensor networks that are best (for different ideas of optimal, i.e. under different cost metrics). However, all these schemes differ from one common failing - they build an optimal tree for a given fixed separation of the sensors. In most practical scenarios, the sensor group is always and dynamically varying - consider for example the set of sensors dotted in a forest that are sensing temperatures above some specified threshold, during a wild fire. Given the limited computational and energy income of sensor nodes it is impossible to either pre store the optimal tree for every possible group or to energetically generate them on the fly.

Lujun Jia, Guevara Noubir, 2011 presented a study on Real-Time Data Aggregation in Contention-Based Wireless Sensor Networks. It examine the problem of delay controlled maximal information collection for CSMA-based wireless sensor networks. They revise how to allocate the maximal acceptable transmission delay at each node, such that the amount of information collected at the sink is maximized and the total delay for the data aggregation is within the given hop. They originate the problem by using lively programming and propose an optimal algorithm for the optimal assignment of transmission attempts. Based on the analysis of the optimal solution, they propose a distributed greedy algorithm. It is shown to have a similar presentation as the best one. CSMA/CA is a contention-based technique.

4. TREEPSI (Tree-based Energy Efficient Protocol for Sensor Information)

4.1 Approach

Multi-hop routing is more preferable for densely deployed networks than making use of single hop communication since multi-hop routing consumes less energy. Some of the assumptions that are made for proposing TREEPSI are illustrated below:

- Nodes are stationary
- Each of its node has the information related to its location
- Sensor network includes energy constrained and homogeneous sensor nodes with uniform initial energy.
• All sensor nodes have power control and have an ability to transmit the information directly to sink or other sensor node.

The first way is to calculate the path centrally through base station and broadcast the information of the path to the network. The second way is that all of the nodes build the same tree structure locally by making use of common algorithm in every node. Once the tree construction is completed, one can make use of any one of the two different ways for gathering data from field.

The two ways are:

• In first method, initially root started data gathering process by sending small control packet to children nodes with the help of standard tree traversal algorithm. Here as the control packet size is minute it consumes very less energy and with little more delay.

• In second method, firstly leaf nodes sense the information and then it just forward this data towards their parent node. Once the parent node receives the data then this node fuse the received information with its own data. The result of this information is forwarded to its parent node. This process continues until the root node receives data.

4.2 Routing Algorithm

In sensor field for each node i do

If (the node is having energy more than what is necessary to transmit a packet to the BS)

then

1) Build a tree structure consisting of all the Live nodes with node i as root node.

2) While (node i is having more than minimum necessary energy for transmission to base station)

do

i) Gather the information in tree structure right from its leaf node to root node as explained in step-1.

ii) Transmit the gathered information to base station by node i.

4.3 Proposed Methodology

In opinion mining are different types of outlook analysis as: feature-level, word level, entity-level, document-level, sentence-level. Data set are collected from different Twitter and micro-blogs by web crawling, in this step will be explained very obviously below paper. Removing data from SMN (Twitter) : using Twitter API REST API s are having the subsequent resources : Time lines , streaming , tweets search, friends, direct message and followers ,users , favorites ,suggested users , lists, saved searchers, Geo , place and trends , spam reports, help. These APIs use the pull tactics for data retrieval. To collect information a user must clearly request it. Streaming APIs provides a nonstop stream of public information from Twitter and micro-blogs. These APIs use the push tactics for data retrieval. Once a demand for information is made, the Streaming APIs provide a continuous torrent of updates with no additional input from the user. Opinion Retrieval involves retrieving preferred information from bag-of-words or Twitter and micro-blogs textual data to measure ad hoc information retrieval usefulness in the standard way; we In many situations, aggregation results received by the BS provide a basis for critical decisions; hence, false or biased aggregation may cause catastrophic consequences. Without loss of generality, it assume
that all the intermediate nodes are aggregators. It claim that the analysis can be easily extended to the case in which only some of the intermediate nodes are aggregators. When a sensor network is organized into a tree topology structure, the aggregation result of every intermediate node is based on its child nodes’ data. To verify that no tampering with the aggregation result at the intermediate node has occurred, let its child nodes do the same aggregation execution and then compare the result derived by the child nodes with the one computed by the parent node. In order to accurately identify the point at which forged aggregation results happen, it limit the commit and-verify scope to every parent–children connection, and verify each intermediate aggregation result. Once there is malicious tampering at any intermediate node, it can immediately find the inconsistency between the committed aggregate and the reconstructed aggregate using LBS level based scheme. In this way, it have accomplished the target of malicious node identification.

4.5 Architecture of Proposed LBS

![Architecture of Proposed LBS](image)

Fig.4.5: Architecture of Proposed LBS

4.6 Algorithm of LBS

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A ix</td>
<td>The row of node i</td>
</tr>
<tr>
<td>Ayi</td>
<td>The column of node i</td>
</tr>
<tr>
<td>Kij</td>
<td>Mediate value of keys</td>
</tr>
<tr>
<td>Tii</td>
<td>Groups</td>
</tr>
<tr>
<td>ri</td>
<td>Random number</td>
</tr>
<tr>
<td>KCi</td>
<td>A key chain maintained by node i</td>
</tr>
</tbody>
</table>

Table 4.2 Notations.

Phase 1: Key Pre-distribution

Step 1: Base Generation

1. The base station generates a large pool of keys (e.g. 520 or more). The keys are selected from a finite filed GF(q) to create a symmetric matrix(SM).

   Where q is the smallest prime larger than the key size.

2. The base station select one publicly known curve K over a finite field eg. F2 as well as to a base point P€K.

Step2: Decompose Matrices to obtain LU Matrices

Base station does the decomposition of the created SM to obtain one lower triangular matrix A and one upper triangular matrix B.

Step3: Key Pre-distribution

Every node is randomly assigned one row from matrix A and one corresponding column from matrix B. For example, node i is assigned row Aix and column Byi, node j is assigned row Ajx and column Byj. After the key pre-distribution, each node only has two vectors in its memory. Each vector has n elements.

Phase 2: Key Establishment

Pairwise Key Establishment Protocol

After key pre-distribution, each node can establish a pairwise key with its neighbors to make sure the secure around communication.

It design a protocol for the process of pairwise key establishment:

1. Node i sends its column Byi to node j.
2. After node j receive Byi, it computes Kji by vector multiplication of Bjx and Byj.

3. Node j reply the Byj, F(Kji) where F(Kji) is the Hash result of the computation of the last step.

4. Upon receiving Byj, node i compute Kij and check if F(Kij) = F(Kji).

5. If it is verified, node i send F(Kij) to node j for the verification.

6. CONCLUSIONS

WSN includes large number of sensor nodes which transfer the data from one system to another system without making use of any wires. All these sensor nodes in the network are resource constrained, so because of this reason the lifetime of the network is limited. Thus, various researchers proposed numerous protocols or approaches for increasing the lifetime of the wireless sensor networks. In this means, the data aggregation concept has been introduced in this report as it is one of the important techniques that enhances the network lifetime. To get a clear idea on all these techniques, a theoretical survey is conducted in this research by referring numerous secondary resources like journals, authorized pdf's, etc. The data was collected by categorizing the research into two parts. In the first part the data was collected on the 82 protocols that are being proposed for data aggregation technique. And in the second part, the data was gathered based on the approaches proposed for securing the WSN data aggregation. In this report, various data aggregation algorithms in WSN are discussed. Further a comprehensive study of different data aggregation protocols are presented under the network architecture. The data aggregation algorithms discussed in this report mainly focuses on three concepts which are efficient routing, organization and data aggregation tree construction. This report described the main features, benefits and limitations of different data aggregation algorithm. However after discussing all the data aggregation protocols it can be concluded that the performance of data aggregation protocol is strongly coupled with network infrastructure.

Even though many of the data aggregation techniques which are discussed looks promising but still there is significant scope for further research. So, this research further extended the work by making study on secure data aggregation since none of the researcher discussed the approaches which provide security to the data aggregation. It's very essential to provide security to the WSN data aggregation since there are numerous effects on network due to lack of security and moreover one cannot increase the lifetime of network without providing security to WSN.

Further of this research clearly discusses the various approaches proposed for securing the data aggregation techniques in WSN. Security can be provided to WSN in various means like false detection, etc. From this study one can easily understand that LBS Level Based Scheme is the most secure to the network and to successfully carry out the data aggregation techniques without Forged Data.

7. REFERENCES

