Cluster Based Data Aggregation for Wireless Sensor Network

SARANYA. V¹, KALAIVANI. A²
Communication systems
Bharathiyar Institute of Engineering for women

Abstract

Data aggregation in wireless sensor networks eliminates redundancy to improve bandwidth utilization and energy efficiency of sensor nodes. One node, called the cluster leader, collects data from surrounding nodes and then sends the summarized information to upstream nodes. This algorithm used to select a cluster leader that will perform data aggregation in a partially connected sensor network. The algorithm reduces the traffic flow inside the network by adaptively selecting the shortest route for packet routing to the cluster leader. It is assumed that an eavesdropping fusion center (EFC) attempts to intercept the transmissions of the sensors and to detect the state of nature. It presents privacy-preserving data aggregation schemes for additive aggregation functions. The scheme is – Cluster-based Private Data Aggregation (CPDA) – leverages clustering protocol and algebraic properties of polynomials. It has the advantage of incurring less communication overhead. The goal of the algorithm is to bridge the gap between collaborative data collection by wireless sensor networks and data privacy. It assesses the scheme by privacy-preservation efficacy, communication overhead, and data aggregation accuracy.

Keywords: Eavesdropping, aggregation, polynomial, preservation

I. INTRODUCTION

A. Wireless Sensor Network

A wireless sensor network (WSN) is an ad-hoc network composed of small sensor nodes deployed in large numbers to sense the physical world. Wireless sensor networks have very broad application prospects including both military and civilian usage. They include surveillance, tracking at critical facilities or monitoring animal habitats. In general, a WSN consists of a large number of tiny sensor nodes distributed over a large area with one or more powerful sinks or base stations (BSs) collecting information from these sensor nodes.

All sensor nodes have limited power supply and have the capabilities of information sensing, data processing and wireless communication. WSN has various characteristics like Ad Hoc deployment, Dynamic network topology, Energy Constrained operation, Shared bandwidth, large scale of deployment. Despite of these characteristics routing in WSN is more challenging. Firstly, resources are greatly constrained in terms of power supply, processing capability and transmission bandwidth. Secondly, it is difficult to design a global addressing scheme as Internet Protocol (IP).

Furthermore, IP cannot be applied to WSNs, since address updating in a large-scale or dynamic WSN can result in heavy overhead. Thirdly, due to the limited resources, it is hard for routing to cope with unpredictable and frequent topology changes, especially in a mobile environment. Fourthly, data collection by many sensor nodes usually results in a high probability of data redundancy, which must be considered by routing protocols. Fifthly, most applications of WSNs require the only communication scheme of many-to-one, i.e., from multiple sources to one particular sink, rather than multicast or peer to peer.

Fig 1.1 Wireless Sensor Network Architecture.
Selecting the optimum sensors and wireless communications link requires knowledge of the application and problem definition.

B. Function Of Data Aggregation

A data aggregation function is taken that aggregates the individual sensor readings. CPDA scheme has focused on additive aggregation function. For computation of the aggregate functions, the following requirements are to be satisfied: (i) privacy of the individual sensor data is to be protected, i.e., each node’s should be known none other expect the node itself, (ii) the number of messages transmitted within the WSN for the purpose of data aggregation should be kept at a minimum, and (iii) the aggregation result should be as accurate as possible. Data aggregation in intermediate nodes (called aggregator nodes) is an effective approach for optimizing consumption of scarce resources like bandwidth and energy in Wireless Sensor Networks (WSNs).

II. RELATED WORK

[1].In this paper the problem of decentralized detection in the presence of one or more classes of misbehaving nodes can be considered. The fusion center first estimates the nodes’ operating points) on the ROC curve and then uses this estimation to classify the nodes and to detect the state of nature. Numerical results are presented that show the proposed algorithm significantly outperforms the reputation-based methods in classification of the nodes as well as the detection of the hypotheses. The estimated operating points are compared to the Cramer- Rao lower bound which shows the efficacy of the proposed method.

[2].In this paper, by exploiting the approximately linear relationship between the scheme parameters and the network size, a simplified q-out-of-m fusion approach for final decision making in the SENMA architecture can be considered. The performance of the proposed scheme is investigated under Byzantine attacks. It was shown that at a fixed percentage of malicious nodes, the false alarm rate of the simplified q-out-of-m scheme decrease exponentially as the network size increase. The detection accuracy of the proposed scheme is further investigated under static and dynamic attacking strategies.

[3].In this paper data confidentiality in a distributed detection scenario with the TBMA protocol in which the wireless channels between the sensors and the ally FC are vulnerable to eavesdropping by an unauthorized enemy FC can be focused. To secure the wireless channels a novel TBMA protocol called secure TBMA which provides data confidentiality by taking advantage of randomness and independence of the main and eavesdropping channels. Instead of securing the individual wireless channels based on cryptographic algorithms, the key idea behind secure TBMA is to have the activated sensors secure their transmissions from possible eavesdropping in a cooperative manner in which the sensors follow different reporting rules depending on the magnitudes of their main channel

[4]. Wireless Sensor Networks often operate in a resource constrained environment. Optimal resource utilization is main objective of WSN. But Wireless Sensor Networks are equally vulnerable to security attacks. Ensuring security in a hostile operational environment of WSN is a hurricane task. The idea of this paper is to provide comprehensive information on types of attacks WSN is exposed to and possible methods of countering such attacks effectively. The motto here is to help novice researchers with objective to work on security challenges in Wireless Sensor Network environment. When wireless sensor networks are deployed in an open or hostile environment security becomes extremely important, as they are prone to different types of malicious attacks.

[5].Based on the canonical parallel fusion structure that incorporates the fading channel, a LR-based fusion rule has been derived. For robust performance in the absence of prior knowledge regarding the local sensors and/or fading channels, several alternatives were proposed. The two stage implementation using the fusion rule provides high SNR approximation to the LR-based fusion rule, whereas the statistic in the form of a MRC statistic gives a low SNR approximation. Performance evaluation is conducted using both the ROC curve as well as the deflection measure. Fusion of binary decisions transmitted over fading channels has particularly important applications in low-cost low-power wireless sensor networks.

III. PROJECT DESCRIPTION

A. Objective
Data aggregation is a process of aggregating the sensor data using aggregation approaches. The general data aggregation algorithm works as shown in the below figure. The algorithm uses the sensor data from the sensor node and then aggregates the data by using some aggregation algorithms such as centralized approach, LEACH, TAG etc.
B. Cluster-Based Approach

In energy-constrained sensor networks of large size, it is inefficient for sensors to transmit the data directly to the sink in such scenarios. Cluster-based approach is hierarchical approach. In cluster-based approach, the whole network is divided into several clusters. Each cluster has a cluster-head which is selected among cluster members. Cluster-heads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to the base station (sink). Recently, several cluster-based network organization and data-aggregation protocols have been proposed for the wireless sensor network.

C. Module Description

1. The Network Model

The basic idea of CPDA is to introduce noise to the raw data sensed from a WSN, such that an aggregator can obtain accurate aggregated information but not individual data points. However, unlike in privacy-preserving data mining where noises are independently generated (at random) and therefore leads to imprecise aggregated results, the noises in CPDA are carefully designed to leverage the cooperation between different sensor nodes, such that the precise aggregated values can be obtained by the aggregator.

2. Key Distribution and Management

CPDA uses a random key distribution mechanism proposed for encrypting messages to prevent message eavesdropping attacks. The key distribution scheme has three phases: (i) key pre-distribution, (ii) shared-key discovery, and (iii) path-key establishment. These phases are described briefly as follows:

These $k$ keys form a key ring for the sensor node. During the key-discovery phase, each sensor node identifies which of its neighbors share a common key with itself by invoking and exchanging discovery messages.

3. Functional Encryption

A closely related notion to homomorphic encryption is functional encryption, where our goal is to reveal the result of the computation to the server, but protect all other information about our encrypted input. For a motivating example consider the problem of spam filtering for encrypted email without interacting with the client. Then the function $f$ we would like to evaluate would be a classifier that sorts e-mails as either "spam" or "not spam". If we had used homomorphic encryption, the server would learn the encrypted bit ($\text{spam/not-spam}$ $\text{Enc}(f(x))$), but it would be useless for sort e-mail messages. Instead like to allow the server to evaluate $\text{Enc}(x) = f(x)$ but just for the particular function $f$ and nothing else (e.g. we don't want the server to compute this for $f(x) = x$).

As a key approach to fulfilling this requirement of private data aggregation, concealed data aggregation schemes have been proposed in which multiple source nodes send encrypted data to a sink along a converge-cast tree with aggregation of cipher-text.

4. Cluster-Based Private Data Aggregation

The CPDA scheme works in three phases: (i) cluster formation, (ii) computation of aggregate results in clusters, and (ii) cluster data aggregation. These phases are described below:

i. Cluster formation

A query server $Q$ triggers a query by a HELLO message. When the HELLO message reaches a sensor node, it elects itself as a cluster leader with a pre-defined probability $p_c$. If a node
becomes a cluster leader, it forwards the HELLO message to its neighbors. If any HELLO message arrives at the node, it decides to join the cluster formed by its neighbor by broadcasting a JOIN message. This process is repeated and multiple clusters are formed so that the entire WSN becomes a collection of a set of clusters.

![Cluster Formation Diagram]

Fig 3.3: Query server Q sends HELLO message to its neighbors. An and D randomly elect themselves as the cluster leaders.

**ii. Computation within clusters:**

In this phase, aggregation is done in each cluster. The computation is illustrated with the example of a simple case where a cluster contains three members: A, B, and C, where A is the assumed as the cluster leader and the aggregator, B and C is the cluster members. Let a, b, c represent the private data held by the nodes A, B, and C respectively.

**iii. Cluster Data Aggregation**

The CPDA scheme has been implemented on top of a protocol known as Tiny Aggregation (TAG) protocol. Using the TAG protocol, each cluster leader node routes the sum of values in its cluster to the query server through a TAG routing tree whose root is situated at the server.

*iv Intermediate node:*

In this module the encrypted data is send to the sink node via intermediate node. This intermediate node calls forward subroutine to forward encrypted data to the sink. The encrypted data is set of cipher texts.

*v. Sink*

Sink is a base station, this is a final stage, and the data are received from different wireless sensor node by this stage. In this stage sink receives the cipher texts and performs decryption process by using the private key. Once the decryption is done plain texts are generated which original message or data is sent by the source (sensor).

**IV. RESULT AND DISCUSSION**

**A. Encryption Function**

Transmitted data can be encrypted and end null packet. The cipher text can be encrypted along with the required random number.

**B. Performance Level Of Encryption**

The various protocol can be reviewed along with number of packet send and also number of receiver can activated for the given transmitter. The authentication rate of various protocol using optimal encryption can be shown in the below figure.

**V. CONCLUSION and FUTURE ENHANCEMENT**

Data privacy is the important parameter for the Secure transmission in wireless sensor network. The optimal probabilistic encryption scheme can be provide more security in the transmission but while considering the data privacy this encryption technique can alter the information. The proposed scheme has no communication overhead and minimal processing requirements making it suitable for sensors with limited resources. In future the clustering based data aggregation scheme will be introduced to transmit with secure.
VI. REFERENCES


