International Journal of Multidisciplinary Research and Development

IJMRD 2015; 2(4): 97-99
www.allsubjectjournal.com
Received: 15-03-2015
Accepted: 02-04-2015
e-ISSN: 2349-4182
p-ISSN: 2349-5979
Impact Factor: 3.762

Prajna Mayadi
M-Tech Scholar,
Department of Cse,
Alvas Institute of
Engineering and
Technology, Moodbidri

Correspondence:
Prajna Mayadi
M-Tech Scholar,
Department of Cse,
Alvas Institute of
Engineering and
Technology, Moodbidri

Data aggregation and privacy preserving techniques in wireless sensor networks: A survey

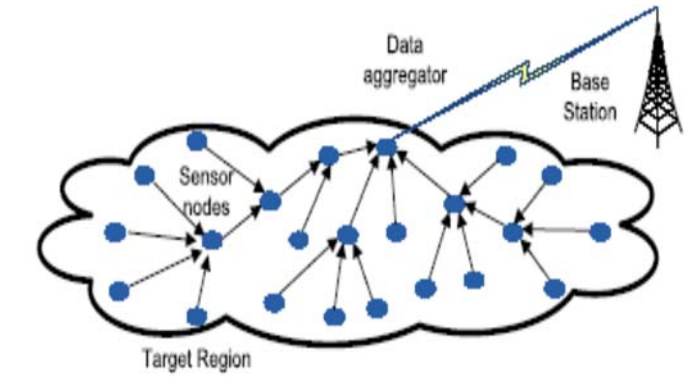
Prajna Mayadi

Abstract
The research advances and applicability of wireless sensor networks (WSNs) have introduced many new promising applications including habitat monitoring, battlefield surveillance and target tracking. The energy consumption is the major issue to be considered for WSNs which is occupied by data communication among nodes in maximum proportion. Many sensor applications collect data from an individual node which is aggregated at a base station. To reduce energy consumption, in-network aggregation can be performed at intermediate nodes en-route to the base station. As wireless sensor networks are usually deployed in remote and hostile environments to transmit sensitive information, sensor nodes are prone to node compromise attacks and security issues such as data confidentiality and integrity are extremely important. Hence, wireless sensor network protocols, e.g., data aggregation protocol, must be designed with security in mind. The paper investigates the relationship between security and data aggregation process.

Keywords: Data Aggregation; Security; Energy efficiency; Wireless Sensor network

1. Introduction
A sensor network consists of a set of battery-powered nodes, which collaborate to perform sensing tasks in a given environment. It may contain one or more base stations to collect sensed data and possibly relay it to a central processing and storage system. Data from sensor nodes are correlated in terms of time and space, transmitting only the required and partially processed data is more meaningful than sending a large amount of raw data. In general, sending raw data wastes energy because duplicated messages are sent to the same node (implosion) and neighboring nodes receive duplicate messages if two nodes share the same observing region (overlap). Thus, data aggregation, which combines data from multiple sensor nodes, has been actively researched in recent years. Since all data are transported wirelessly between sensor nodes, they are typically prone to interception and eavesdropping. Speaking broadly, there are two types of privacy concerns in WSNs: internal privacy and external privacy. The former is about maintaining the data privacy of a sensor node from other trusted participating sensor nodes of the WSN, whereas the latter means that the sensed data is protected from outsiders (adversaries). Data privacy can be simply defined as a process in which private data can be overheard and decrypted by adversaries or other trusted participating sensor nodes, but it can still provide a mechanism that prevents them from recovering sensitive information, *i.e.*, control disclosure of any information about the data. To achieve data privacy, it is required to protect transmission trend of a node's private data from its neighboring nodes. This is because the neighboring nodes can always overhear the sum of the private data and a fixed unknown number, *i.e.*, an encryption key.

2. Data Aggregation in Wsn
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 Fig. 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 multihop path.



2.1 Approaches to data aggregation

Tree-Based Approach [1]: In the tree-based approach perform aggregation by constructing an aggregation tree, which could be a minimum spanning tree, rooted at sink and source nodes are considered as leaves. Each node has a parent node to forward its data. Flow of data starts from leaves nodes up to the sink and therein the aggregation done by parent nodes.

Cluster-Based Approach [2]: In cluster-based approach, whole network is divided in to several clusters. Each cluster has a cluster-head which is selected among cluster members. Clusterheads do the role of aggregator which aggregate data received from cluster members locally and then transmit the result to sink.

3. Wireless Sensor Networking Requirements and Challenges

For a wireless sensor network to deliver real-world benefits, it must support the following requirements in deployment: scalability, reliability, responsiveness, mobility, and power efficiency. The complex inter-relationships between these characteristics is a balance; if they are not managed well, the network can suffer from overhead that negates its applicability in the real world. In order to ensure that the network supports the application's requirements, It is important to understand how each of the wireless sensor networking characteristics affects reliability.

Table 1: Essential Requirements of Wireless Sensor Networks

Requirements	Description
Reliability	The ability of the network to ensure reliable data transmission in a state of continuous change of network structure.
Scalability	The ability of the network to grow, in terms of the number of nodes, without excessive overhead.
Responsiveness	The ability of the network to quickly adapt itself to changes in topology.
Mobility	The ability of the network to handle mobile nodes and changeable data paths.
Power efficiency	The ability of the network to operate at extremely low power levels.

4. Routing

Multihop routing is a critical service required for WSN. Because of this, there has been a large amount of work on this topic. Internet and MANET routing techniques do not perform well in WSN. Internet routing assumes highly reliable wired connections so packet errors are rare; this is not true in WSN. Many MANET routing solutions depend on symmetric links (i.e., if node A can reliably reach node B, then B can reach A) Between neighbours; this is too often not true for WSN. These differences have necessitated the invention and deployment of new solutions. For WSN, which are often deployed in an ad hoc fashion, routing typically begins with neighbour discovery. Nodes send rounds of messages (packets) and build local neighbour tables. These tables include the minimum information of each neighbour's ID and location. This means that nodes must know their geographic location prior to neighbour discovery. Other typical information in these tables includes nodes' remaining energy, delay via that node, and an estimate of link quality. Once the tables exist, in most WSN routing algorithms messages are directed from a source location to a destination address based on geographic coordinates, not IDs. A typical routing algorithm that works like this is Geographic Forwarding (GF) .In GF, a node is aware of its location, and a message that it is "routing" contains the destination address. This node can then computes which neighbour node makes the most progress towards the destination by using the distance formula from geometry. It then forwards the message to this next hop. In variants of GF, a node could also take into account delays, reliability of the link and remaining energy.

Another important routing paradigm for WSN is directed diffusion [11]. This solution integrates routing,

Queries and data aggregation. Here a query is disseminated indicating an interest in data from remote nodes. A node with the appropriate requested data responds with an attribute-value pair. This attribute- value pair is drawn towards the requestor based on gradients, which are set up and updated during query

dissemination and response. Along the path from the source to the destination, data can be aggregated to reduce communication costs. Data may also travel over multiple paths increasing the robustness of routing.

5. Privacy preserving data aggregation schemes

Homomorphic Scheme is used to design two privacy preserving data aggregation schemes for additive aggregation functions: CPDA and SMART. In both the schemes the information from a sensor node is known only to that sensor node only while others will get the aggregated values of the readings.

CPDA: He *et al.* proposed the Cluster-based Private Data Aggregation (CPDA) [3] to achieve privacy-preserving data aggregation for WSNs. In the CPDA, sensor nodes are randomly grouped into clusters for creating an aggregation tree. Each cluster leverages the additive property of polynomials to calculate the desired aggregate value. At the same time, it guarantees that no individual node can know the data values of other nodes. The intermediate aggregate values in each cluster will be further aggregated on the way to the data sink along the aggregation tree. First, every sensor node in each cluster customizes its private data into polynomial form of order $k - 1$, where k is the total number of nodes in a cluster using shared (non-private) seeds and random numbers (private). Secondly, each sensor node encrypts its customized value by using a unique shared key between a sensor node and the other sensor nodes of the cluster. Thirdly, all nodes from the same cluster exchange their encrypted customized data with each other. Each sensor node has to encrypt and decrypt $O(Nc)$ messages, where Nc is the number of sensor nodes in a cluster. Each node assembles all the data including its own by using the additive property of polynomials and sends them to their respective cluster leaders. After that, the cluster leaders

deduce the aggregate value by computing the inverse of an $M \times M$ matrix where M is the number of cluster nodes. Finally, each cluster leader routes the derived sum of the cluster back towards the query server through the TAG routing tree [4].

SMART: The Slice-Mix-Aggregate (SMART) by He *et al.* [3] achieves privacy-preserving data aggregation by hiding original data before the data transmissions. For this, each sensor node first customizes its private data by slicing it into a fixed number of pieces. Then, it sends data slices to a particular number of neighboring sensor nodes. After the data pieces are received from the neighboring sensor nodes, all sensor nodes calculate the aggregate value of the data slices so that the privacy of the sensor data can be preserved. In the SMART, each sensor node randomly selects a set of sensor nodes, say J, within h hops. When each sensor node slices its private data randomly into J pieces, $J - 1$ pieces are encrypted and sent to the randomly selected sensor nodes, keeping one data piece at the same sensor node. All the sensor nodes decrypt the data by using their shared keys and sum all the received slices. Each sensor node sends the sum to its parent. Finally, the root of the network is the ultimate aggregation point of all sensor data

References

1. M. Lee and V.W.S. Wong, “An Energy-aware Spanning Tree Algorithm for Data Aggregation in Wireless Sensor Networks,” IEEE PacRrim 2005,Victoria, BC, Canada, Aug. 2005.
2. K. Dasgupta, K. Kalpakis, and P. Namjoshi, “An Efficient Clustering-based Heuristic for Data Gathering And Aggregation in Sensor Networks”, IEEE 2003.
3. He, W.; Liu, X.; Nguyen, H.; Nahrstedt, K.; Abdelzaher, T. Pda: Privacy-preserving data aggregation in wireless sensor networks. In Proceeding of the 26th IEEE International Conference on Computer Communications, Anchorage, AK, USA, May 6–12, 2007; pp. 2045–2053..
4. Madden, S.R.; Franklin, M.J.; Hellerstein, J.M.; Hong, W. TAG: A tiny aggregation service for ad hoc sensor networks. In Proceedings of the 5th Symposium on Operating Systems Design and Implementation, Boston, MA, USA, December 9–11, 2002; pp. 1–16.