International Journal of Computing and Corporate Research
ISSN (Online) : 2249-054X
Volume 4 Issue 3 May 2014
International Manuscript ID : 2249054XV4I3052014-02

AN EFFICIENT ALGORITHMIC APPROACH FOR DATA AGGREGATION IN WIRELESS SENSOR NETWORKS

Jamal kh-Madhloom al-bdairi (M.Tech)
CSE Department
M. M. Engineering College,
M. M. University,
Mullana, Ambala, Haryana, India

Rohit Vaid
CSE Department
M. M. Engineering College,
M. M. University,
Mullana, Ambala, Haryana, India

Abstract

Data aggregation in Wireless Sensor Network refers to acquiring the sensed data from the sensors to the gateway node. Data aggregation plays a vital role in Wireless Sensor Networks since the aggregation schemes followed here involve in reducing the amount of power consumed during data transmission between the sensor nodes. There are several aggregation techniques followed in Wireless Sensor Network, such as Tree based Aggregation and in network aggregation. A Wireless Sensor Network (WSN) typically consists of a sink node sometimes referred to as a Base Station and a number of small wireless sensor nodes. The base station is assumed to be secure with unlimited available energy while the sensor nodes are assumed to be unsecured with limited available energy. The sensor nodes monitor a geographical area and collect sensory information. Sensory information is communicated to the Base Station through Wireless hop by hop transmissions. To conserve energy this information is aggregated at intermediate sensor nodes by applying a suitable aggregation function on the received data. While traditionally encryption is used to provide end to end confidentiality in Wireless Sensor Network (WSN), the aggregators in a secure data aggregation scenario need to decrypt the encrypted data to perform aggregation. This exposes the plaintext at the aggregators, making the data vulnerable to attacks from an adversary. Similarly an aggregator can inject false data into the aggregate and make the base station accept false data. Thus, while data aggregation improves energy efficiency of a network, it complicates the existing security challenges. In this manuscript, the data aggregation factors are analyzed in the wireless sensor networks.

Keywords - Data Aggregation, WSN, Wireless Sensor Networks

International Journal of Computing and Corporate Research
ISSN (Online) : 2249-054X
Volume 4 Issue 3 May 2014
International Manuscript ID : 2249054XV4I3052014-02

1. Introduction

The recent advances in Micro Electro Mechanical System (MEMS) technology, low cost and low power consumption, wireless micro sensor nodes have been available. Wireless sensor networks (WSN) usually consist of a base station and many sensor nodes. The sensor nodes are randomly distributed over the sensor network's field. The sensor nodes monitor environmental factors such astemperature, air pressure, and motion, and then send this sensing data to the base station. The base station acts as a gateway to deliver information from the sensor nodes to outside users who need it. In WSN, it is too difficult to initialize the sensor nodes and manage the sensor networks due to the large number of sensor nodes, which may number tens of thousands. Therefore, self-configuring sensor nodes are desirable in WSN. Moreover, in order to save energy, sensor nodes carry out data aggregation and perform compression on the gathered data before sending data to the base station, and execute energy efficient routing [1]. Sensor nodes can send their data to the base station by direct communication protocol or a multi-hop communication method such as the Minimum Transmission Energy (MTE) routing protocol. In direct communication protocol, sensor nodes, which are far from the base station, dissipate faster than others do because they send their data to the base station directly; sensor nodes do not only transmit their own sensing data, but also serve as routers for other sensor nodes if they use the MTE routing protocol. Therefore, the energy of the sensor nodes that are near to the base station is rapidly consumed in the MTE routing protocol [4].

2. Clustering in Wireless Sensor Networks

In most wireless sensor network (WSN) applications nowadays the entire network musthave the ability to operate unattended in harsh environments in which pure human access and monitoring cannot be easily scheduled or efficiently managed or it's even not feasible at all.

Based on this critical expectation, in many significant WSNapplications the sensor nodes are often deployed randomly in the area of interest byrelatively uncontrolled means (i.e., dropped by a helicopter) and they form a network in an ad hoc manner. Moreover, considering the entire area that has to be covered, the short duration of the battery energy of the sensors and the possibility of havingdamaged nodes during deployment, large populations of sensors are expected; it's anatural possibility that hundreds or even thousands of sensor nodes will be involved. Inaddition, sensors in such environments are energy constrained and their batteries usually cannot be recharged. Therefore, it's obvious that specialized energy-aware routing and data gathering protocols offering high scalability should be applied in order that network lifetime is preserved acceptably high in such environments. Naturally, grouping sensor nodes into clusters has been widely adopted by the research community to satisfy the above scalability objective and generally achieve high energy efficiency and prolong network lifetime in large-scale WSN environments. The corresponding hierarchical routing and data gathering protocols imply cluster-based organization of the sensor nodes in order that data fusion and aggregation are possible, thusleading to significant energy savings.

3. Main Objectives and Design Challenges ofClustering in WSNs

As was mentioned at the beginning, hierarchical clustering in WSNs can greatlycontribute to overall system scalability, lifetime, and energy efficiency.Hierarchicalrouting is an efficient way to lower energy consumption within a cluster, performingdata aggregation and fusion in order decrease the number of transmitted messages to theBS.On the contrary, a single-tier network can cause the gateway to overload with theincrease in sensors density. Such overload might cause latency in communication and inadequate tracking of events. In addition, the single-tier architecture is not scalable for a larger set of sensors covering a wider area of interest because the sensors are typically not capable of long-haul communication. Hierarchical clustering is particularly useful for applications that require scalability to hundreds or thousands of nodes. Scalability in thiscontext implies the need for load balancing and efficient resource utilization. Applications requiring efficient data aggregation (e.g., computing the maximum detected radiation around a large area) are also natural candidates for clustering. Routing protocols can also employ clustering. Clustering was also proposed as a useful tool for efficiently pinpointing object locations. In addition to supporting network scalability and decreasing energy consumption through data aggregation, clustering has numerous other secondary advantages and corresponding objectives.

4. Classification of Clustering Algorithms

a) Low Energy Adaptive Clustering Hierarchy (LEACH)

One of the first and most popular clustering protocols proposed for WSNs was LEACH (Low Energy Adaptive Clustering Hierarchy). It is probably the first dynamic clustering protocol which addressed specifically the WSNs needs, using homogeneous stationary sensor nodes randomly deployed, and it still serves as the basis for other improved clustering protocols for WSNs.

b) Energy-Efficient Hierarchical Clustering (EEHC)

Another significant probabilistic clustering algorithm was earlier proposed (Energy Efficient Hierarchical Clustering - EEHC). The main objective of this algorithm was to address the shortcomings of one-hop random selection algorithms such as LEACH by extending the cluster architecture to multiple hops.

c) Hybrid Energy-Efficient Distributed Clustering (HEED)

Another improved and very popular energy-efficient protocol is HEED (Hybrid Energy- Efficient Distributed Clustering). HEED is a hierarchical, distributed, clustering scheme in which a single-hop communication pattern is retained within each cluster, whereasmulti-hop communication is allowed among CHs and the BS.

d) Weight-Based Clustering Protocols

In addition to node proximity, some other known algorithms use a combination of metrics such as the remaining energy, transmission power, etc., (thus forming corresponding combined weights) to achieve more generalized goals than single-criterion protocols. Several algorithms following this directive were initially borrowed from the field of mobile ad hoc networks.

5. Requirement of Clustering:

a) To identify the taxonomy of the training sets
b) To analyze the properties of the set
c) The perform specific operations based on the taxonomy
d) To investigate the grouping and classifications

6. Data Aggregation and Techniques

Data aggregation is a type of data and information mining process where data is searched, gathered and presented in a report-based, summarized format to achieve specific business objectives or processes and/or conduct human analysis.Data aggregation may be performed manually or through specialized tool.

Data aggregation in Wireless Sensor Network refers to acquiring the sensed data from the sensors to the gateway node.Data aggregation plays a vital role in Wireless Sensor Networks since the aggregation schemes followed here involve in reducing the amount of power consumed during data transmission between the sensor nodes.
There are several aggregation techniques followed in Wireless Sensor Network such as

- Tree based Aggregation
- In network aggregation

A Wireless Sensor Network (WSN) typically consists of a sink node sometimes referred to as the sensor nodes monitor a geographical area and collect sensory information. Sensory information is communicated to the Base Station through Wireless hop by hop transmissions. To conserve energy this information is aggregated at intermediate sensor nodes by applying a suitable aggregation function on the received data. Aggregation reduces the amount of network traffic which helps to reduce energy consumption on sensor nodes. It however complicates the already existing security challenges for wireless sensor networks and requires new security techniques tailored specifically for this scenario. Providing security to aggregate data in Wireless Sensor Networks is known as Secure Data Aggregation in WSN. were the first few works discussing techniques for secure data aggregation in Wireless Sensor Networks.

Two main security challenges in secure data aggregation are confidentiality and integrity of data. While traditionally encryption is used to provide end to end confidentiality in Wireless Sensor Network (WSN), the aggregators in a secure data aggregation scenario need to decrypt the encrypted data to perform aggregation. This exposes the plaintext at the aggregators, making the data vulnerable to attacks from an adversary. Similarly an aggregator can inject false data into the aggregate and make the base station accept false data. Thus, while data aggregation improves energy efficiency of a network, it complicates the existing security challenges.

7. Review of Literature

Novel Cluster Based Routing Protocol- This work has proposed a novel Cluster Based Routing Protocol (CBRP) for prolong the sensor network lifetime. CBRP achieves a good performance in terms of lifetime by balancing the energy load among all the nodes. In this protocol first we Cluster the network by using new factors and then construct a spanning tree for sending aggregated data to the base station which can better handle the heterogeneous energy capacities.

An unequal cluster-based routing protocol- Clustering provides an effective method for prolonging the lifetime of a wireless sensor network. Current clustering algorithms usually utilize two techniques; selecting cluster heads with more residual energy, and rotating cluster heads periodically to distribute the energy consumption among nodes in each cluster and extend the network lifetime. However, they rarely consider the hot spot problem in multi hop sensor networks. When cluster heads cooperate with each other to forward their data to the base station, the cluster heads closer to the base station are burdened with heavier relay traffic and tend to die much faster, leaving areas of the network uncovered and causing network partitions. To mitigate the hot spot problem, they've proposes an Unequal Cluster-based Routing (UCR) protocol. It groups the nodes into clusters of unequal sizes. Cluster heads closer to the base station have smaller cluster sizes than those farther from the base station, thus they can preserve some energy for the inter-cluster data forwarding. A greedy geographic and energy-aware routing protocol is designed for the inter-cluster communication, which considers the tradeoff between the energy cost of relay paths and the residual energy of relay nodes.

SajidHussain and Abdul W. MatinJodrey School of Computer Science, Acadia University Wolfville, Nova Scotia, Canada–The Hierarchical cluster-based routing (HCR) technique is an extension of the LEACH [1] protocol that is a selforganized cluster-based approach for continuous monitoring. In LEACH, the network is randomly divided into several clusters, where each cluster is managed by a cluster head (CH). The sensor nodes transmit data to their cluster heads, which transmit the aggregated data to the base station. In HCR, each cluster is managed by a set of associates and the energy efficient clusters are retained for a longer period of time; the energy-efficient clusters are identified using heuristics-based approach. Moreover, in a variation of HCR, the base station determines the cluster formation. A Genetic Algorithm (GA) is used to generate energy-efficient hierarchical clusters. The base station broadcasts the GA-based clusters configuration, which is received by the sensor nodes and the network is configured accordingly**.**

An Enhanced Cluster Based Routing Algorithm - The efficient node-energy utilization is one of important performance factors in wireless sensor networks because sensor nodes operate with limited battery power. They've proposed a cluster based routing algorithm to extend the lifetime of the networks and to maintain a balanced energy consumption of nodes. To obtain it, the authors add a tiny slot in a round frame, which enables to exchange the residual energy messages between the base station (BS), cluster heads, and nodes. The slot is used in the Pre-setup phase. The performance of the proposed protocol has been examined and evaluated with the NS-2 simulator. As a result of simulation, they have confirmed that our proposed algorithm shows the better performance in terms of lifetime than LEACH.A cluster-based routing protocol for wireless sensor networks with nonuniform node distribution

JiguoYu,YingyingQi,GuanghuiWang,XinGu, School of Computer Science, Qufu Normal University, Rizhao, Shandong, China -The energy consumption among nodes is more imbalanced in cluster-based wireless sensor networks. Based on this problem, in that paper, a cluster-based routing protocol for wireless sensor networks with non uniform node distribution is proposed, which includes an energy-aware clustering algorithm EADC and a cluster-based routing algorithm. EADC uses competition range to construct clusters of even sizes. At the same time, the routing algorithm increases forwarding tasks of the nodes in scarcely covered areas by forcing cluster heads to choose nodes with higher energy and fewer member nodes as their next hops, and finally, achieves load balance among cluster heads.

International Journal of Computing and Corporate Research
ISSN (Online) : 2249-054X
Volume 4 Issue 3 May 2014
International Manuscript ID : 2249054XV4I3052014-02

Clustering-based power-controlled routing– The authors have presented two new routing protocols for mobile sensor networks, viz. power-controlled routing (PCR) and its enhanced version, i.e. Enhanced Power-Controlled Routing (EPCR). In both the protocols, fixed transmission power is employed in the clustering phase but when ordinary nodes are about to send their data to their respective cluster-heads, they change their transmission power according to their distance from their cluster-head. While in PCR, the nodes are associated with the cluster-head on the basis of weight, in EPCR it is done on the basis of distance. In addition to the protocols, the authors suggest a packet loss recovery mechanism for the PCR and EPCR. Both protocols work well for both mobile and static networks and are designed to achieve high network lifetime, high packet delivery ratio, and high network throughput. These protocols are extensively simulated using mass mobility model, with different speeds and different number of nodes to evaluate their performance.

Some of the steps followed by existing techniques are
Step 1: Cluster Heads broadcasts identifier message to all other wireless sensor nodes (Adjacent Nodes). The adjacent nodes replies with identifier acknowledgement to the cluster head.
Step 2: The Cluster Head manages the routing table and also the details of all the nodes in its group. The Cluster Head also maintains details about other groups Cluster Head and its address with the help of Base station.
Step 3: The normal sensor node in a group maintains a table that contains information of its Cluster Head address and the common identifier generated by the Cluster Head.
Step 4: The address of the Cluster Head that has already involved in routing has stored in every packet, it is used for verification by other Cluster Head.
Step 5: When a source node in a need of route to deliver packets to the destination node, it sends Route Request message to the Cluster Head, the Cluster Head uses its common identifier to verify the packets.
Step 6: The Cluster Head checks whether the destination node is in house, if the destination node is present under its group, then it sends the packet directly. If the destination node is not in house then it sends Route Request message to Base Station, The Base Station intent passes it to the Cluster Head which manages the Destination node. The Cluster Head passes then passes packets to the Destination node.
Step 7: The sensor node under motion makes new route request to the Cluster Head, then Cluster Head passes the information to the base station. Source node and destination node under range directly communicates with each other with the help of adjacent nodes.
Step 8: Suppose the link with the cluster head fails then the packet transfer process will also fail.

8. Comparison of properties of different Data Aggregation Approaches

Model	**Sensor Failure Resilience**	**Confidentiality**	**Message Length independent of WSN Size**
CMT	No	Yes	No
HBH	Yes	No	Yes
No Aggregation	Yes	Yes	No

Scheme	**Confidentiality**	**Integrity**	**Freshness**	**Authentication**
CDA	**Y**			
SDA		**Y**		**Y**
SIA	**Y**	**Y**	**Y**	**Y**
SHDA		**Y**	**Y**	**Y**
WDA		**Y**	**Y**	
SECUREDAV	**Y**	**Y**		**Y**
SRDA	**Y**	**Y**	**Y**	**Y**
SDAP	**Y**	**Y**	**Y**	**Y**
ESA	**Y**	**Y**	**Y**	**Y**
EDA	**Y**	**Y**		**Y**

9. Summary of Network Flow Based Data Aggregation Algorithms

Algorithm	**Objective**	**Approach**	**Limitations**
CMLDA	Network Lifetime maximize	Integer Linear Programming	High Complexity
MFA	Network Lifetime maximize	Dijkstra Shortest Path Tree Algorithm	Performance decrease
RFEC Algorithm	High Performance	Graph Based Approach	Performance
SPTA	Min. Cost	Joint Optimization	Correlation

10. CONCLUSION

Aggregation reduces the amount of network traffic which helps to reduce energy consumption on sensor nodes. It however complicates the already existing security challenges for wireless sensor networks and requires new security techniques tailored specifically for this scenario. Providing security to aggregate data in Wireless Sensor Networks is known as Secure Data Aggregation in WSN were the first few works discussing techniques for secure data aggregation in Wireless Sensor Networks. Two main security challenges in secure data aggregation are confidentiality and integrity of data.

References

[1]Novel Cluster Based Routing Protocol in Wireless Sensor Networks by BagerZarei, Mohammad Zeynali and VahidMajidNezhad Department of Computer Engineering, Islamic Azad University, Shabestar Branch Tabriz, East-Azarbaijan, Iran.

[2]An unequal cluster-based routing protocol in wirelesssensor networks byGuihai Chen · Chengfa Li · Mao Ye · JieWu.Department of Computer Science and Engineering, Florida Atlantic University Boca Raton, FL 33431, USA.

[3]SajidHussain and Abdul W. MatinJodrey School of Computer Science, Acadia University Wolfville, Nova Scotia, Canada.

[4]An Enhanced Cluster Based Routing Algorithm for Wireless Sensor Networks by Uk-Pyo Han, Sang-Eon Park, Seung-Nam Kim, Young-Jun Chung Computer Science Department, Kangwon National University, Chunchon, Korea.

[5]A cluster-based routing protocol for wireless sensor networks with nonuniform node distributionJiguoYu,YingyingQi,GuanghuiWang,Xin Gu, School of Computer Science, Qufu Normal University, Ri-zhao, Shandong 276826, PR China.

[6]Clustering-based power-controlled routing for mobile wireless sensor networks by Atta urRehman Khan, Sajjad A. Madani,Khizar Hayat, SameeUllah Khan.