Computer Networks 53 (2009) 2022–2037

Contents lists available at ScienceDirect

Computer Networks

journal homepage: www.elsevier.com/locate/comnet

# Secure data aggregation in wireless sensor networks: A comprehensive overview ☆

Suat Ozdemir [a,*], Yang Xiao [b]

[a] Computer Engineering Department, Gazi University, Maltepe, Ankara, TR-06570, Turkey
[b] Department of Computer Science, The University of Alabama, Tuscaloosa, AL 35487-0290, United States

ARTICLE INFO

*Article history:*
Received 8 July 2008
Received in revised form 29 November 2008
Accepted 22 February 2009
Available online 24 March 2009

Responsible Editor: J. Misic

*Keywords:*
Secure data aggregation
Survey
Wireless sensor networks

ABSTRACT

Wireless sensor networks often consists of a large number of low-cost sensor nodes that have strictly limited sensing, computation, and communication capabilities. Due to resource restricted sensor nodes, it is important to minimize the amount of data transmission so that the average sensor lifetime and the overall bandwidth utilization are improved. Data aggregation is the process of summarizing and combining sensor data in order to reduce the amount of data transmission in the network. 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. This paper investigates the relationship between security and data aggregation process in wireless sensor networks. A taxonomy of secure data aggregation protocols is given by surveying the current "state-of-the-art" work in this area. In addition, based on the existing research, the open research areas and future research directions in secure data aggregation concept are provided.

## 1. Introduction

Wireless sensor networks are usually composed of hundreds or thousands of inexpensive, low-powered sensing devices with limited memory, computational, and communication resources [1,2]. 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 [6]. Hence, it is a challenging task to provide efficient solutions to data gathering problem. Among these constraints, "battery power" is the most limiting factor in designing wireless sensor network protocols. Therefore, in order to reduce the power consumption of wireless sensor networks, several mechanisms are proposed such as radio scheduling, control packet elimination, topology control, and most importantly data aggregation [2,3]. 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 neighboring nodes, aggregates them (e.g., computes the average), and sends the aggregated data to the base station

☆ Dr. Ozdemir's work is supported in parts by the Gazi University Scientific Research Project Fund No. 06/2007-44. Prof. Y. Xiao's work is supported in parts by the US National Science Foundation (NSF) under Grants Nos. CCF-0829827, CNS-0716211, and CNS-0737325.

* Corresponding author. Tel.: +90 312 582 3123; fax: +90 312 230 8434.
*E-mail addresses:* suatozdemir@gazi.edu.tr, suatozdemir@hotmail.com (S. Ozdemir), yangxiao@cs.ua.edu (Y. Xiao).

1389-1286/$ - see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.comnet.2009.02.023

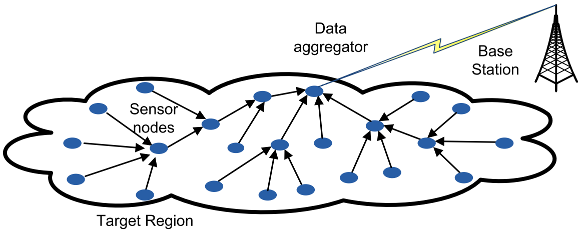

**Fig. 1.** Data aggregation in a wireless sensor network.

over a multihop path. As illustrated by the example, data aggregation reduces the number of data transmissions thereby improving the bandwidth and energy utilization in the network.

In wireless sensor networks, the benefit of data aggregation increases if the intermediate sensor nodes perform data aggregation incrementally when data are being forwarded to the base station. However, while this continuous data aggregation operation improves the bandwidth and energy utilization, it may negatively affect other performance metrics such as delay, accuracy, fault-tolerance, and security [3]. As the majority of wireless sensor network applications require a certain level of security, it is not possible to sacrifice security for data aggregation. In addition, there is a strong conflict between security and data aggregation protocols. Security protocols require sensor nodes to encrypt and authenticate any sensed data prior to its transmission and prefer data to be decrypted by the base station [26,29]. On the other hand, data aggregation protocols prefer plain data to implement data aggregation at every intermediate node so that energy efficiency is maximized. Moreover, data aggregation results in alterations in sensor data and therefore it is a challenging task to provide source and data authentication along with data aggregation. Due to these conflicting goals, data aggregation and security protocols must be designed together so that data aggregation can be performed without sacrificing security.

The necessity of implementing data aggregation and security together have led many researchers to work on secure data aggregation problem. In this paper, we aim to provide an extensive overview of secure data aggregation concept in wireless sensor networks by defining the main issues and covering the most important work in the area. Compared to general data aggregation problem which is a well researched topic in wireless sensor networks, secure data aggregation problem still has the potential to provide many interesting research opportunities. Hence, we also aim to give a starting point for researchers who are interested in secure data aggregation problem by presenting the open research areas and future research directions in the field.

Our contribution in this paper is twofold. First, we look at the data aggregation problem from the security perspective by giving a comprehensive literature survey. Second, based on the observations from the state-of-the-art secure data aggregation protocols, we discuss the open research areas and future research directions. Although there are couple of existing survey papers on data aggregation in wireless sensor networks [7,8], to the best of our knowledge, this is the first survey paper that focuses on solely secure data aggregation concept. In this paper, we cover many secure data aggregation protocols that are not covered by the previous survey papers. In addition, the open research areas and future research directions presented in this paper do not appear in the existing survey papers either. Nonetheless, we believe our paper will serve as a useful guide and starting point for the researchers who are interested in conducting research in the secure data aggregation area. The organization of the paper as follows: Section 2 starts with a brief summary of security requirements of wireless sensor networks and show how they relate with data aggregation process. Section 3 gives introductory information about data aggregation and summarize the most important work in the area. Section 4 presents “state-of-the-art” secure data aggregation protocols in wireless sensor networks. In this section, a broad overview of secure data aggregation is given by evaluating each protocol based on the security requirements of wireless sensor networks. Section 5 defines open research areas and future research directions in secure data aggregation. Section 6 concludes the paper by emphasizing our contributions in this paper.

## 2. Security requirements of wireless sensor networks

Due to hostile environments and unique properties of wireless sensor networks, it is a challenging task to protect sensitive information transmitted by wireless sensor networks [1]. In addition, wireless sensor networks have security problems that traditional networks do not face. Therefore, security is an important issue for wireless sensor networks and there are many security considerations that should be investigated. In this section, we present the essential security requirements that are raised in a wireless sensor network environment and explain how these requirements relate with data aggregation process. Fig. 2 illustrates the interaction between wireless sensor network security requirements and data aggregation process.

### *2.1. Data confidentiality*

In wireless sensor networks, data confidentiality ensures that secrecy of sensed data is never disclosed to unauthorized parties and it is the most important issue in mission critical applications. Authors of [4] state that a sensor node should not leak its readings to neighboring nodes. Moreover, in many applications, sensor nodes transmit highly sensitive data, e.g., secret keys; and therefore it is extremely important to build secure channels among sensor nodes. Public sensor information, such as sensor identities and public keys, should also be encrypted to some extent to protect against traffic analysis attacks. Furthermore, routing information must also remain confidential in certain cases as malicious nodes can use this information to degrade the network's performance. The standard approach for keeping sensitive data secret is to encrypt the data with a secret key that only intended receivers possess, hence achieving confidentiality. However, data aggregation protocols usually cannot aggregate

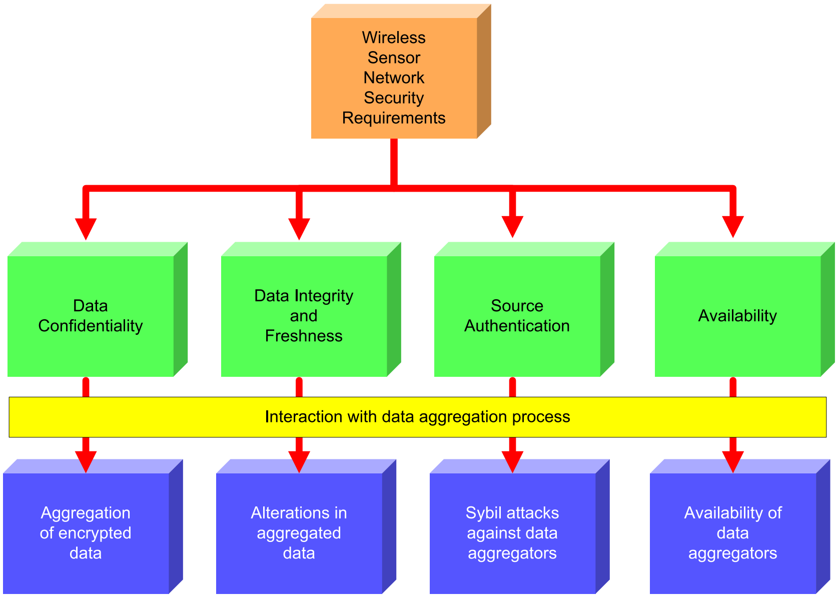
Wireless Sensor Network Security Requirements
Data Confidentiality
Data Integrity and Freshness
Source Authentication
Availability
Interaction with data aggregation process
Aggregation of encrypted data
Alterations in aggregated data
Sybil attacks against data aggregators
Availability of data aggregators

**Fig. 2.** Interaction between wireless sensor network security and data aggregation process.

encrypted data. Therefore, such data aggregation protocols must decrypt the sensor data to perform data aggregation and encrypt the aggregated data before transmitting it. This decryption/encryption of sensor data at data aggregators not only results in delay and energy consumption but also prevents end-to-end data confidentiality.

### 2.2. Data integrity and freshness

Although data confidentiality guarantees that only intended parties obtain the un-encrypted plain data, it does not protect data from being altered. Data integrity guarantees that a message being transferred is never corrupted. A malicious node may just corrupt messages to prevent network from functioning properly. In fact, due to unreliable communication channels, data may be altered without the presence of an intruder. Thus, message authentication codes or cyclic codes are used to prevent data integrity. Data aggregation results in alterations of data; therefore, it is not possible to have end-to-end integrity check when data aggregation is employed. Moreover, if a data aggregator is compromised, then it may corrupt sensor data during data aggregation and the base station has no way of checking the integrity of this aggregated sensor data. Providing data integrity is not enough for wireless communication because compromised sensor nodes are able to listen to transmitted messages and replay them later on to disrupt the data aggregation results. Data freshness protects data aggregation schemes against replay attacks by ensuring that the transmitted data is recent.

### 2.3. Source authentication

Since wireless sensor networks use a shared wireless medium, sensor nodes need authentication mechanisms to detect maliciously injected or spoofed packets. Source authentication enables a sensor node to ensure the identity of the peer node it is communicating with. Without source authentication, an adversary could masquerade a node, thus gaining unauthorized access to resource and sensitive information and interfering with the operation of other nodes. Moreover, a compromised node may send data to its data aggregator under several fake identities so that the integrity of the aggregated data is corrupted. Faking multiple sensor node identities is called Sybil attack and it poses significant threat to data aggregation protocols [5]. If only two nodes are communicating, authentication can be provided by symmetric key cryptography. The sender and the receiver share a secret key to compute the message authentication code (MAC) for all transmitted data. However, data aggregators may need broadcast authentication which requires more complex techniques, such as $\mu$TESLA [5].

### 2.4. Availability

Availability guarantees the survivability of network services against Denial-of-Service (DoS) attacks. A DoS attack can be launched at any layer of a wireless sensor network and may disable the victim node(s) permanently. In addition to DoS attacks, excessive communication or computation may exhaust battery charge of a sensor node. Consequences of availability loss may be catastrophic. For example, in a battlefield surveillance application, if the availability of some sensor nodes cannot be provided, this may lead to an enemy invasion. Wireless sensor networks are deployed with high node redundancy to tolerate such availability losses. Since data aggregators collect the data of a number of sensor nodes and sends the aggregated data to the base station, availability of data aggregators is more important than regular sensor nodes. Thus, in wireless sensor networks, intruders launch DoS attacks with the aim of preventing data aggregators from performing their task so that some part of the network losses its availability.

## 3. Data aggregation

In a typical wireless sensor network, a large number of sensor nodes collect application specific information from the environment and this information is transferred to a central base station where it is processed, analyzed, and used by the application. In these resource constrained networks, the general approach is to jointly process the data generated by different sensor nodes while being forwarded toward the base station [8]. Such distributed in-network processing of data is generally referred as *data aggregation* and involves combining the data that belong the same phenomenon. The main objective of data aggregation is to increase the network lifetime by reducing the resource consumption of sensor nodes (such as battery energy and bandwidth). While increasing network lifetime, data aggregation protocols may degrade important quality of service metrics in wireless sensor networks, such as *data accuracy*, *latency*, *fault-tolerance*, and *security*. Therefore, the design of an efficient data aggregation protocol is an inherently challenging task because the protocol designer must trade off between energy efficiency, data accuracy, latency, fault-tolerance, and security. In order to achieve this trade off, data aggregation techniques are tightly coupled with how packets are routed through the network. Hence, the architecture of the sensor network plays a vital role in the performance of different data aggregation protocols. There are several protocols that allow routing and aggregation of data packets simultaneously. These protocols can be categorized into two parts: *tree-based* data aggregation protocols and *cluster-based* data aggregation protocols. Earlier work on data aggregation focused on improving the existing routing algorithms so as to make data aggregation possible. As a result, many data aggregation protocols based on shortest path tree structure have been proposed [10,17,46]. To reduce the latency due to tree-based data aggregation, recent work on data aggregation tends to group sensor nodes into clusters so that data are aggregated in each group for improved efficiency.

### *3.1. Tree-based data aggregation protocols*

The simplest way to achieve distributed data aggregation is to determine some data aggregator nodes in the network and ensure that the data paths of sensor nodes include these data aggregator nodes. Such tree-based data aggregation techniques have been extensively studied in the literature [9–18]. The main issue of tree-based data aggregation protocols is the construction of an energy efficient data aggregation tree. Fig. 3 illustrates an example of tree-based data aggregation. Greedy Incremental Tree (GIT) [9] is a data-centric routing protocol that allows data aggregation based on Directed Diffusion [10]. In [11] GIT is compared with two other data-centric routing schemes, namely Center at Nearest Source (CNS) [3] and Shortest Path Tree (SPT) [10]. The simulation results show that GIT performs the best in terms of average number of transmissions. Another SPT based data aggregation protocol that promotes the parent energy-awareness is proposed in [12]. In this protocol, parent selection is based on sensor nodes' distance to the base station and their residual energy level. There are also data aggregation protocols that consider information theory as routing metric. For example, [13] proposes a centralized approach that routes the packet based on their joint entropies. However, this protocol is not feasible as it depends on the global knowledge of the information entropy of each sensor node as well as the joint entropy of each node pair. In the rest of this subsection, we present some of the important work in tree-based data aggregation in detail.

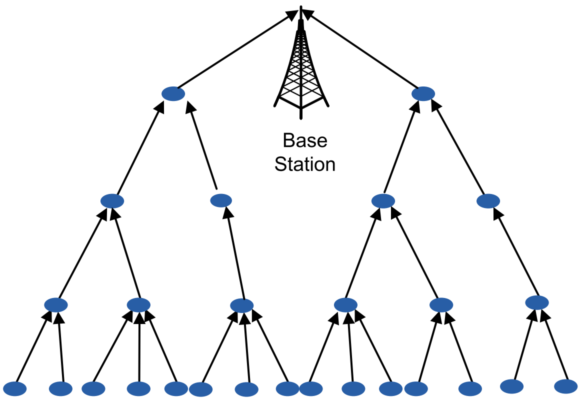

**Fig. 3.** Tree-based data aggregation.

In [14], Madden et al. proposed a data-centric data aggregation framework called Tiny AGgregation Service (TAG), which is based on shortest path tree routing. TAG is specifically designed for monitoring applications and allows an adjustable sleep schedule for sensor nodes. To achieve this, parent nodes let their children know about the waiting time for transmission. Also, parent nodes cache their children's data to prevent from data loss. TAG performs data aggregation in two phases. In the first phase, called distribution phase, base station queries are disseminated to the sensor nodes, then in the second phase, called collection phase, the aggregated sensor readings are routed up the aggregation tree. During the distribution phase, a message is broadcasted by the base station requiring sensor nodes to organize a routing tree so that the base station can send its queries. Each message has a field that specifies the level or distance from the root of the sending node (the level of the root is equal to zero). When a node that does not belong to any level receives this message, it sets its own level by incrementing the current level in the message by one and assigns the sender as its parent. This process continues until all sensor nodes in the network joins the tree and have a parent. This messaging periodically repeated to keep the tree structure updated. Once the tree is formed, then the base station queries the network via the aggregation tree. Sensor nodes use their parents when replying to base station queries. TAG employs an SQL like language to query the network. Each query specifies the quantity that needs to be collected, aggregation function and the sensor nodes that need to perform the data collection.

Directed diffusion [10] is a reactive data-centric protocol which takes places in three phases (i) interest dissemination (ii) gradient setup, and (iii) path reinforcement and forwarding. In the first phase, the base station propagates an interest message (interest dissemination) that describes the type of data that needs to be collected and the operational mode for the collection. Upon reception of this message, each sensor node rebroadcasts the message to its neighbors. Sensor nodes also prepare interest gradients which are basically the vectors containing the next hop that has to be used to propagate the result of the query back to the base station (gradient setup). For each type of data a different gradient may be set up. At end of gradient setup phase for a certain type of data, only a single path is used to route packets toward the sink (path reinforcement and forwarding). An illustrative example of directed diffusion protocol is presented in Fig. 4. Data aggregation is performed during data forwarding phase. The base station periodically refreshes the data gathering tree which is formed by the reinforced paths. However, this is an expensive operation and it may overcome the gain by the data aggregation if the network has a dynamic topology. A modified version of directed diffusion, called Enhanced Directed Diffusion (EDD), is proposed in [15] which integrates directed diffusion with a cluster-based architecture so that the efficiency of the local interactions during gradient set up phase increases. Another similar protocol is proposed in [16].

Power-Efficient GAthering in Sensor Information Systems (PEGASIS) that organizes sensor nodes in a chain for the purpose of data aggregation is proposed in [17]. In PEGASIS, each data aggregation chain has a leader that is responsible to transmit aggregated data to the base station. In order to evenly distribute the energy expenditure in the network, sensor nodes take turns acting as the chain leader. The chain forming can be achieved either in centralized manner by the base station or in a decentralized manner by using a greedy algorithm at each sensor node. Both approaches require the global knowledge of the network. The chain building process starts from the sensor node furthest from the base station and continues towards the base station. When a node dies, the chain is reconstructed to bypass the dead node. In a sensor node chain, each sensor node receives data from a neighbor and aggregates it with its own reading by generating a single packet that has the same length with the received data. This process is repeated along the chain and the leader adds its own data into the packet and sends it to the base station directly. It should be noted that node *i* will be in some random position *j* on the chain. Thus, the leader in each round of communication will be at a random position on the chain, which is important to make the sensor network robust to node failures. Two major drawbacks of PEGASIS have been observed. First, PEGASIS requires each sensor node to have a complete view of the network topology so that chains can be formed properly. In addition, all nodes must be able to transmit directly to the base station. Second, if the distances between sensor nodes in a chain are too big, then the energy expenditure of sensor nodes can be significantly high.

A data aggregation tree construction protocol that only relies on local knowledge of the network topology is proposed in [12]. The proposed protocol, called EADAT (Energy-Aware Distributed Aggregation Tree), is based on an energy-aware distributed heuristic. The base station is the root of the aggregation tree hence it initiates the tree forming by broadcasting a control message which has the following five fields: ID, parent, power, status, and hopcount. This message is forwarded among sensor nodes until each node broadcasts the message once and the result is an aggregation tree rooted at the base station. By considering energy level of sensor nodes, the algorithm gives higher chance to sensor nodes with higher residual power to become a non-leaf tree node. Therefore, data forwarding task is performed by the sensor nodes that have high energy levels. Simulation results show that EADAT prolongs network lifetime and saves more energy in comparison with routing methods without aggregation. It is also observed that the average energy level of sensor nodes decreases much more slowly compared to the scenario without data aggregation.

There are many additional solutions that solve the problem of efficiently constructing data aggregation trees in wireless sensor networks. A different approach, called Delay Bounded Medium Access Control (DBMAC), that integrates routing and MAC protocols to perform data aggregation is proposed in [18]. The main objective of the proposed DBMAC scheme is both to minimize the latency for delay bounded applications and to increase energy efficiency by taking advantage of data aggregation mechanisms. DBMAC employs a carrier sense multiple access with collision avoidance (CSMA/CA) medium accesses scheme based on a request to send/clear to send/data/acknowledgment (RTS/CTS/DATA/ACK) handshake. By taking advantage of CTS messages of other nodes, sensor nodes can select the relay node among those nodes that already have some packets to transmit in their queue. This process increases the data aggregation efficiency in the network as all the information stored along the path is aggregated into a singe data packet. DBMAC is an excellent

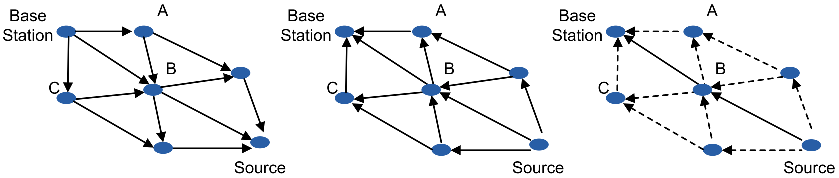

**Fig. 4.** Illustrative example of directed diffusion. If the base station sends an interest that reaches sensor nodes *A* and *B*, and both forward the interest to sensor node *C*, then node *C* sets up two vectors indicating that the data matching the interest must be returned to *A* and/or *B*.

example of how routing and data aggregation may influence each other by showing that energy efficient data aggregation solutions are obtained by a cross-layer design.

### 3.2. Cluster-based data aggregation protocols

In cluster-based data aggregation protocols, sensor nodes are subdivided into clusters. In each cluster, a cluster head is elected in order to aggregate data locally and transmit the aggregation result to the base station. Cluster heads can communicate with the sink directly via long range radio transmission. However, this is quite inefficient for energy constrained sensor nodes. Thus, cluster heads usually form a tree structure to transmit aggregated data by multihopping through other cluster heads which results in significant energy savings. Fig. 5 presents an example of cluster-based data aggregation. Recently, several cluster-based data aggregation protocols have been proposed [19–25].

In [19], a self-organizing and adaptive clustering protocol, called Low-Energy Adaptive Clustering Hierarchy (LEACH) is proposed. LEACH takes advantage of randomization to evenly distribute the energy expenditure among the sensor nodes. LEACH is a clustered approach where cluster heads act as data aggregation points. The protocol consists of two phases. In the first phase, cluster structures are formed. Then, in the second phase, cluster heads aggregate and transmit the data to the base station. LEACH's cluster head election process is based on a distributed probabilistic approach as follows. In each data aggregator selection round, sensor nodes calculate the threshold $T(n)$:

$$T(n) = \begin{cases} \frac{P}{1-P(R\,\mathrm{mod}(1/P))} & \text{if } n \in G, \\ 0 & \text{otherwise}. \end{cases}$$

Here $P$ is the desired percentage of cluster heads, $R$ is the round number, and $G$ is the set of nodes that have not been cluster heads during the last $1/P$ rounds. In order to be a cluster head, a sensor node $n$ picks a random number between $[0,1]$ and becomes a cluster head if this number is lower than $T(n)$. Cluster head advertisements are broadcasted to sensor nodes and sensor nodes join the clusters based on the signal strength of the advertisement messages. Based on the number of cluster members, each cluster head schedules its cluster-based on TDMA to optimally manage the local transmissions. In the second phase, sensor nodes send their data to cluster heads according to the established schedule. Optionally, sensor nodes may turn off their radios until their scheduled TDMA transmission slot. LEACH requires cluster heads to send their aggregated data to the base station over a single link. However, this is a disadvantage of LEACH because single link transmission may be quite expensive when the base station is far away from the cluster head. LEACH is completely distributed as it does not require any global knowledge regarding network structure. It is also an adaptive protocol in terms of cluster head selection. On the other hand, there may be high control message overhead if the network topology is dynamic due to mobil nodes.

Another cluster-based data aggregation protocol, called HEED, is proposed in [20]. For cluster head selection, HEED benefits from the availability of multiple power levels at sensor nodes. In fact, a combined metric that is composed of the node's residual energy and the node's proximity to its neighbors. HEED defines the average of the minimum power level required by all sensor nodes within the cluster to reach the cluster head. This is called Average Minimum Reachability Power (AMRP). AMPR is used to estimate the communication cost in each cluster. In order to select cluster heads, each sensor node computes its probability of becoming the cluster head as follows:

$$P_{(CH)} = C \times \frac{E_{residual}}{E_{max}},$$

where $C$ and $E_{residual}$ and $E_{max}$ denote the initial percentage of cluster heads, the current residual, and initial energy of the sensor node, respectively. Each sensor node broadcasts a cluster head message, sensor nodes select their cluster head as the node with the lowest AMRP in the set of received cluster head messages. This process recursively continues until every node is assigned to a cluster head. As in LEACH, cluster heads in HEED, communicate directly with the base station. Simulation results show that HEED extends the network lifetime and results in geographically balanced set of cluster heads.

In [21], the authors propose a clustering scheme that performs periodic per hop data aggregation. The proposed scheme is called Cougar and it is suitable for applications where sensor nodes continuously generate correlated data.

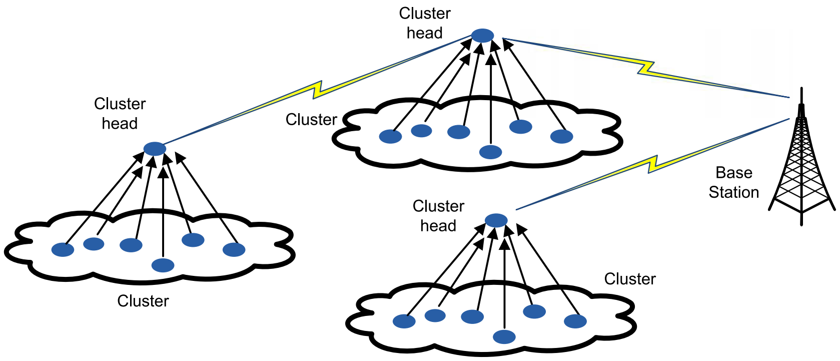

**Fig. 5.** Cluster-based data aggregation.

Once cluster heads aggregate their cluster data, they send the local aggregated data to a gateway node. Similar to LEACH, Cougar is negatively affected by dynamic network topologies. However Cougar has a unique cluster head election procedure. Cougar selects the cluster heads based on more than one metric and allows sensor nodes to be more than one hop away from their cluster heads. This calls for routing algorithms to exchange packets within clusters. Cougar employs the Ad Hoc On Demand Distance Vector (AODV) protocol for inter cluster relaying. In Cougar, synchronization is used to correctly aggregate the data. The cluster head is synchronized with all sensor nodes in the cluster and it does not report its aggregated data to the gateway node until all sensor nodes send their data. Therefore, the synchronization mechanism help improving the correctness of the aggregated data.

Clustered Diffusion with Dynamic Data Aggregation (CLUDDA) [22] is a hybrid approach that combines clustering with diffusion mechanisms. CLUDDA includes query definitions inside interest messages which are initiated by the base station. Each interest message contains the definition of the query that describes the operations that need to be performed on the data components in order to generate a proper response. Interest transformation reduces the processing overhead by utilizing the existing knowledge of queries. CLUDDA combines directed diffusion [10] and clustering during the initial phase of interest propagation. Using clustering mechanism, it is ensured that only cluster heads that perform inter cluster communication are involved in the transmission of interest messages. As the regular sensor nodes do not transmit any data unless they are capable of servicing a request, CLUDDA conserves energy. In CLUDDA, any cluster head that has the knowledge of query definition can perform data aggregation, and hence the aggregation points are dynamic. Also, each cluster head maintains a query cache to present the different data components that were aggregated to obtain the final data. Cluster heads also keep a list of the addresses of neighboring nodes from which the data messages originated. These addresses are used to propagate interest messages directly to specific nodes instead of broadcasting.

There are other cluster-based data aggregation algorithms in the literature. Some of these are improvements of existing protocols. In [24], a cross-layer approach is adopted by integrating medium access control scheme design into a data aggregation concept. A location-based clustering scheme where the sensor nodes self-organize to form static clusters is proposed in [25]. In this protocol, sensor nodes send their data to cluster head along shortest paths, and in-network aggregation is performed at the intermediate nodes. Cluster heads perform aggregation and send aggregated data to the base station over multihop paths. However, during aggregated data transmission from cluster heads to the base station no further aggregation is performed.

## 4. Secure data aggregation

Like any other wireless sensor network protocol, data aggregation protocols must satisfy the security requirements explained in Section 2. However, the resource constrained sensor nodes and necessity of plain data for aggregation process pose great challenges when implementing security and data aggregation together. Security requirements of wireless sensor networks can be satisfied using either *symmetric key* or *asymmetric key* cryptography. Due to resource constraints of sensor nodes, symmetric key cryptography is preferable over asymmetric key cryptography. Hence, the necessity of implementing data aggregation and security using symmetric key cryptography algorithms have led many researchers to work on *secure data aggregation* problem [26–34]. In these protocols, security and data aggregation are achieved together in a hop-by-hop fashion. That is, data aggregators must decrypt every message they receive, aggregate the messages according to the corresponding aggregation function, and encrypt the aggregation result before forwarding it. In addition, these schemes require data aggregators to establish secret keys with their neighboring nodes. Therefore, hop-by-hop secure data aggregation protocols cannot provide data confidentiality at data aggregators and result in latency because of the decryption/encryption process. In order to mitigate the drawbacks of hop-by-hop secure data aggregation protocols, a set of data aggregation protocols is proposed [36–41]. The proposed protocols perform data aggregation without requiring the decryption of the sensor data at data aggregators. While some of these protocols use symmetric cryptography, others employ asymmetric key cryptography functions, such as [42,43], that are suitable for resource constrained sensor nodes. As data aggregators do not have to decrypt sensor data to perform aggregation, the protocols proposed in [36–41] provide end-to-end data confidentiality and result in less latency compared to hop-by-hop secure data aggregation protocols. On the other hand, the downside of the data aggregation protocols that do not require the decryption of sensor data is that they are applicable to only a set of aggregation functions, such as sum and average. In what follows, we classify and explain the secure data aggregation protocols based on the requirement of decrypting sensor data at data aggregators.

### 4.1. Secure data aggregation using plain sensor data

Earlier work on secure data aggregation is focused on symmetric key cryptography and aggregation of plain data. In [26], the authors propose security mechanisms to detect node misbehavior (dropping, modifying or forging messages, transmitting false aggregate value). The key idea of this work is delayed aggregation. Instead of aggregating messages at the immediate next hop, messages are forwarded unchanged over the first hop and then aggregated at the second hop. This is achieved using a key chain, the base station periodically broadcast authentication keys. Hence, sensor nodes need to buffer the data to authenticate it once the authentication key is broadcasted by the base station. The proposed protocol ensures data integrity, and however it does not provide data confidentiality. In addition, if a parent node and its child are compromised nodes, then data integrity is not guaranteed either.

In [27], random sampling mechanisms and interactive proofs are used to check the correctness of the aggregated data at the base station. The proposed protocol is called SIA. The authors claim that, by constructing efficient random sampling mechanisms and interactive proofs, it is possible for the user to verify that the aggregated data provided by the aggregator is a good approximation of the true value even when the aggregator and a fraction of the sensor nodes are compromised. In particular, the authors present efficient protocols for securely computing the median and the average of the measurements, estimating of the network size, and finding the minimum and maximum sensor reading. In the paper, the correctness of data is checked by constructing a Merkle hash tree. In this construction, all the collected data is placed at the leaves of the tree, and the aggregator computes a binary hash tree starting from the leaf nodes: each internal node in the hash tree is computed as the hash value of the concatenation of the two child nodes. The root of the tree is called the commitment of the collected data. Fig. 6 shows an example of Merkle hash tree construction. The hash function in use has to be collision resistant. Once the aggregator commits to the collected values by sending those values to base station, it cannot change any of the collected values. The authors in [27] also assume that each sensor node has a unique identifier and shares a separate secret cryptographic key with the base station and with the aggregator. These keys enable data confidentiality, integrity and authentication.

SecureDAV protocol [28] is very similar to [27] except that elliptic curve cryptography is used for encryption purposes. Moreover, SecureDAV improves the data integrity vulnerability by signing the aggregated data. SecureDAV is a clustered approach where all sensor nodes within a cluster share a secret cluster key. Each sensor node is able to generate a partial signature over the aggregated data. Each data aggregator aggregates its cluster data and broadcasts the aggregated data to its cluster. Each sensor node in the cluster compares its data with the aggregated data broadcasted by the data aggregator. A sensor node partially signs the aggregated data if and only if the difference between its data and aggregated data is less than a threshold. Finally, the data aggregator combines the partial signatures to form a full signature of the aggregated data and sends it to the base station. SecureDAV provides data confidentiality, data integrity, and source authentication. However, the scheme incurs high communication overhead on data validation and supports only the average aggregation function.

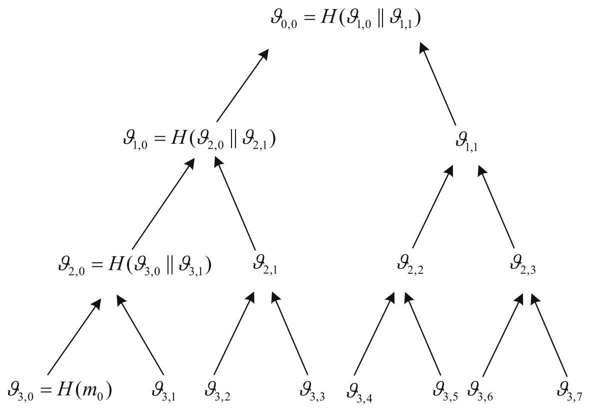
$\vartheta_{0,0} = H(\vartheta_{1,0} \parallel \vartheta_{1,1})$
$\vartheta_{1,0} = H(\vartheta_{2,0} \parallel \vartheta_{2,1})$
$\vartheta_{1,1}$
$\vartheta_{2,0} = H(\vartheta_{3,0} \parallel \vartheta_{3,1})$
$\vartheta_{2,1}$
$\vartheta_{2,2}$
$\vartheta_{2,3}$
$\vartheta_{3,0} = H(m_0)$
$\vartheta_{3,1}$
$\vartheta_{3,2}$
$\vartheta_{3,3}$
$\vartheta_{3,4}$
$\vartheta_{3,5}$
$\vartheta_{3,6}$
$\vartheta_{3,7}$

**Fig. 6.** An example of Merkle hash tree construction to commit to a set of values. $H(m_0)$ represents hash of message $m_0$. Each $\vartheta$ is result of its hash of its two children values. If any node of tree is changed, $\vartheta$ value at the root changes.

A witness based data aggregation scheme for wireless sensor networks is proposed in [30]. The witness nodes of each data aggregator also perform data aggregation and compute MACs of the aggregated data. Witness nodes do not send their aggregated data to the base station. Instead, each witness node sends its MAC of the aggregated data to the data aggregator. The data aggregator collects and forwards the MACs to the base station. Those MACs that are computed by the witness nodes are used at the base station for verifying the correctness of the data aggregated by data aggregators. This enhances the assurance of data aggregation. In order to prove the validity of the aggregated data, each data aggregator has to provide proofs from several witnesses. Because the data validation is performed at the base station, the transmission of false data and MACs up to base station affects adversely the utilization of sensor network resources. The proposed protocol offers only integrity property to the data aggregation security.

In [31], sensor nodes use the cryptographic algorithms only when a cheating activity is detected. Topological constraints are introduced to build a secure aggregation tree (SAT) that facilitates the monitoring of data aggregators. In SAT, any child node is able to listen to the incoming data of its parent node. When the aggregated data of a data aggregator are questionable, a weighted voting scheme is employed to decide whether the data aggregator is properly behaving or is cheating. If the data aggregator is a misbehaving node, then SAT is rebuilt locally so that the misbehaving data aggregator is excluded from the aggregation tree.

In [33], a Secure Hop-by-hop Data Aggregation Protocol (SDAP) is proposed. The authors of SDAP are motivated by the fact that, compared to low-level sensor nodes, more trust is placed on the high-level nodes (i.e., nodes closer to the root) during a normal hop-by-hop aggregation process in a tree topology. Because aggregated data calculated by a high-level node represents the data of a large number of low-level sensor nodes. If a compromised node is closer to the base station, the false aggregated data produced by this compromised node will have a larger impact on the final result computed by the base station. Since all sensor nodes have simple hardware that is prone to compromise, none of those low-cost sensor nodes should be more trustable than others. Hence, SDAP aims to reduce the approach of reducing the trust on high-level nodes by following the divide-and-conquer principle. SDAP dynamically partitions the topology tree into multiple logical groups (subtrees) of similar sizes using a probabilistic approach. In this way, fewer nodes are located under a high-level sensor node in a logical subtree resulting in reduced potential security threat by a compromised high-level node. An example of a grouped tree is shown in Fig. 7. SDAP provides data confidentiality, source authentication, and data integrity.

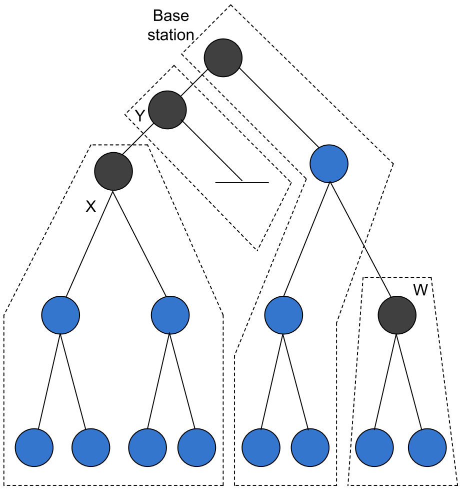

**Fig. 7.** An example of the aggregation tree in SDAP. The nodes *X*, *Y* and *W* with the color black are leader nodes, and the base station as the root is a default leader.

In [34], the authors argue that compromised nodes have access to cryptographic keys that are used to secure the aggregation process and therefore cryptographic primitives alone cannot provide a sufficient enough solution to secure data aggregation problem. Based on this observation, the authors propose a *S*ecure and r*EL*iable *D*ata *A*ggregation protocol, called SELDA which makes use of a *web of trust*. The basic idea behind SELDA is that sensor nodes observe actions of their neighboring nodes to develop trust levels (trustworthiness) for both the environment and the neighboring nodes. As shown in Fig. 8, sensor nodes employ monitoring mechanisms to detect node availability, sensing and routing, misbehaviors of their neighbors. These misbehaviors are quantified as trust levels using Beta distribution function [44,45]. Sensor nodes exchange their trust levels with neighboring nodes to form a web of trust that allows them to determine secure and reliable paths to data aggregators. Moreover, to improve the reliability of the aggregated data, data aggregators weigh sensor data they receive using the web of trust. One important property of SELDA is that, due to the monitoring mechanisms in use, it can detect if a data aggregator is under DoS attack. The simulation results show that SELDA increases the reliability of the aggregated data at the expense of a tolerable communication overhead. In [35], the authors improved the main idea of SELDA by introducing functional reputation concept where each functional reputation value is computed over sensor node actions with respect to that function. Hence, security of data aggregation process is ensured by selecting trusted data aggregators using *aggregation* functional reputation and by weighting sensor data using *sensing* functional reputation. The simulation results show that using functional reputation is more effective than using general reputation when evaluating the trustworthiness of a sensor node.

In wireless sensor networks, a compromised sensor node can inject false data during data forwarding and aggregation to forge the integrity of aggregated data. It is highly desirable for sensor nodes to detect and drop false data as soon as possible in order to avoid depleting their limited resources such as battery power and bandwidth [49]. Although several secure data aggregation protocols [27,28,30] are able to detect the false data injected by sensor nodes, false data injections by compromised data aggregators cannot be detected by these methods. The reason is that data aggregation results in data alterations and therefore a change in aggregated data due to false data injection is extremely hard to detect. Such false data injections by compromised data aggregators can easily result in false alarms that waste the network's resources and reduce the operational efficiency [49]. Recently, some work has been reported on detection false data injections during data aggregation so that the false alarm ratio in the network is reduced [47–49].

In [47,48] secure data aggregation problem is addressed from intrusion detection perspective. In the proposed scheme, an Extended Kalman Filter (EKF) based mechanism to detect false injected data is proposed. Along with the employment of EKF, the proposed mechanism monitors sensor nodes to predict their future real in-network aggregated values. For aggregated values, a normal range is determined to detect false data injections. Using different aggregation functions (average, sum, max, and min), the authors show how to obtain normal ranges theoretically. Moreover, it is also shown that the proposed EKF is

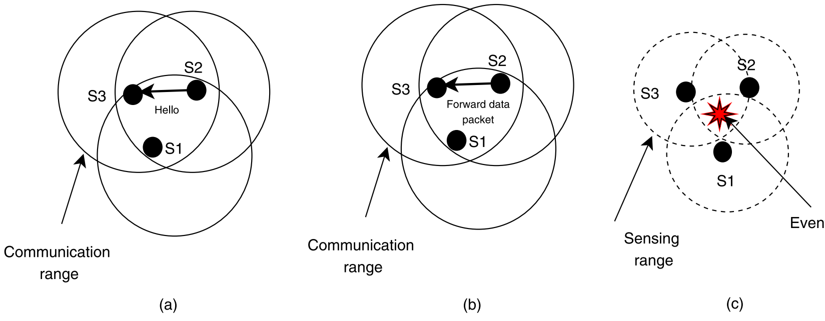

**Fig. 8.** (a) S1 detects node availability misbehavior of S3, if S3 does not respond S2's Hello messages over a period of time. (b) S1 detects routing misbehavior of S3, if S3 does not forward S2's data packets properly. (c) Event E is detected by S1, S2 and S3, if event E is reported falsely by any one of these nodes, the sensing misbehavior of that node is detected by the other two nodes.

used to create effective local detection mechanisms. The created local detection approaches are able to differentiate between malicious events and emergency events and therefore it can reduce the false alarm rate in the network. Extensive simulations are performed to evaluate performance of local detection mechanisms, including false positive rate and detection rate, under different aggregation functions. Simulation results demonstrate that the proposed techniques achieve desirable performance to detect false injected data.

The work presented in [49] realizes the fact that many existing false data detection techniques consider false data injections during data forwarding only. The paper presents a data aggregation and authentication protocol, called DAA, to integrate false data detection with data aggregation and confidentiality. To support data aggregation along with false data detection, a monitoring algorithm is proposed. Using this monitoring algorithm, the monitoring nodes of every data aggregator also conduct data aggregation and compute the corresponding small-size message authentication codes for data verification at their pairmates. To support confidential data transmission, the sensor nodes between two consecutive data aggregators verify the data integrity on the encrypted data rather than the plain data. Each data packet is appended with two full-size message authentication codes, each consisting of $T+1$ small-size message authentication codes. Performance analysis shows that DAA detects any false data injected by up to $T$ compromised nodes, and that the detected false data are not forwarded beyond the next data aggregator on the path. Despite that false data detection and data confidentiality increase the communication overhead, simulation results show that DAA can still reduce the amount of transmitted data by up to 60% with the help of data aggregation and early detection and dropping of false data.

The authors of [50] address how to determine a false alarm threshold dynamically and efficiently in order to minimize the false alarm probability in a wireless sensor networks deployed in realistic environments. In the proposed dynamic threshold scheme, the threshold changes in accordance with the false alarm rate. Hence, a better detection probability and reduced number of false alarms are achieved. Considering the realistic deployment scenarios, the paper proposes to reduce the impact of noise by taking a weighted average of different sensing units readings for the same target. The paper takes advantage of the fact that sensing units of different types are affected at varying degrees by the environmental factors. The authors also propose a data aggregation algorithm to determine the detection probability of a target by fusing data from multiple sensors. Although data confidentiality and authentication are not considered in the proposed data aggregation algorithm, the simulation results show that it improves the target detection accuracy and minimize false alarm rate in the network.

All of the above secure data protocols use actual sensor data for aggregation and hence require decryption of sensor data at aggregators. However, the protocols proposed in [29,32] do not need actual data and therefore they are able to integrate security and data aggregation seamlessly. In [29], the authors present an energy efficient and Secure Pattern based Data Aggregation (ESPDA) protocol which considers both data aggregation and security concepts together in cluster-based wireless sensor networks. ESPDA is the first protocol to consider data aggregation techniques without compromising security. ESPDA uses pattern codes to perform data aggregation. The pattern codes are representative data items that are extracted from the actual data in such a way that every pattern code has certain characteristics of the corresponding actual data. The extraction process may vary depending on the type of the actual data. For example, when the actual data are images of human beings sensed by the surveillance sensors, the key parameter values for the face and body recognition are considered as the representative data depending on the application requirements. When a sensor node consists of multiple sensing units, the pattern codes of the sensor node are obtained by combining the pattern codes of the individual sensing units. Instead of transmitting the whole sensed data, sensor nodes first generate and then send the pattern codes to cluster heads. Cluster heads determine the distinct pattern codes and then request only one sensor node to send the actual data for each distinct pattern code. This approach makes ESPDA both energy and bandwidth efficient. ESPDA is also secure because cluster heads do not need to decrypt the data for data aggregation and no encryption/decryption key is broadcast. Additionally, the proposed nonblocking OVSF (Orthogonal Variable Spreading Factor) block hopping technique further improves the security of ESPDA by randomly changing the mapping of data blocks to NOVSF time slots.

In [32], Secure Reference-Based Data Aggregation (SRDA) protocol is proposed for cluster-based wireless sensor networks. Like ESPDA, SRDA also realizes the fact that data aggregation protocols should work in conjunction with the data communication security protocols, and that any conflict between these protocols might create loopholes in-network security such as violating data confidentiality. In SRDA, raw data sensed by sensor nodes are compared with reference data values and then only the difference data are transmitted. Reference data is taken as the average value of a number of previous sensor readings. The motivation behind SRDA is that it is critical to reduce the number of bits in a transmission because radio communication is the most energy-consuming activity in a sensor node. While data aggregation reduces the number of packets, decreasing the *size* of the transmitted packets will further improve the energy savings. In conventional data aggregation algorithms, sensors transmit their raw sensed data to the cluster heads. This wastes energy and bandwidth since a certain range of the data may remain the same in each packet. However, SRDA transmits the differential data rather than the raw sensed data. That is, the raw data sensed by sensor nodes are compared with reference data and then only the difference data are transmitted. As an example, let 102°F denote the temperature measurement of a sensor node. If 100°F is considered as reference temperature by the cluster head, the sensor node can send only the difference (i.e., 2°F) of the current measurement from the reference value in the transmission. Consequently, differential aggregation has great potential to reduce the amount of data to be transmitted from sensor nodes to cluster heads. The downside of ESPDA [29] and SRDA [32] is that they do not allow intermediate nodes to perform data aggregation. That is, sensor data can be aggregated only at the immediate data aggregator which significantly limits the benefit of data aggregation. We present the data aggregation protocols that do not require decryption of sensor data but also allow intermediate nodes to perform data aggregation in the next subsection.

### 4.2. Secure data aggregation using encrypted sensor data

By using traditional symmetric key cryptography algorithms, it is not possible to achieve end-to-end confidentiality and in-network data aggregation together. If the application of symmetric key based cryptography algorithms is combined with the requirement of efficient data aggregation, then the messages must be encrypted hop-by-hop. However, this means that, in order to perform data aggregation, intermediate nodes have to decrypt each received message, then aggregate the messages according to the corresponding aggregation function, and finally encrypt the aggregation result before forwarding it. Clearly, this is not an energy efficient way of performing secure data aggregation and it may result in considerable delay. In addition, this process requires neighboring data aggregators to share secret keys for decryption and encryption. In order to achieve end-to-end data confidentiality and data aggregation together without requiring secret key sharing among data aggregators privacy homomorphic cryptography has been used in the literature [36–40].

A privacy homomorphism is an encryption transformation that allows direct computation on encrypted data. Let $E$ denotes *encryption* and $D$ denotes *decryption*. Also let $+$ denotes addition and $\times$ denotes multiplication operation over a data set $Q$. Assume that $K_{pr}$ and $K_{pu}$ are the private and public keys of the base station, respectively. An encryption transformation is accepted to be additively homomorphic, if

$$a + b = D_{K_{pr}}(E_{K_{pu}}(a) + E_{K_{pu}}(b)) \quad \text{where } a, b \in Q$$

and it is accepted to be multiplicatively homomorphic, if

$$a \times b = D_{K_{pr}}(E_{K_{pu}}(a) \times E_{K_{pu}}(b)) \quad \text{where } a, b \in Q$$

Since, additively and multiplicatively homomorphic cryptographic functions support additive and multiplicative operations on encrypted data, respectively, data aggregators can perform addition and multiplication based data aggregation over the encrypted data.

In Concealed Data Aggregation (CDA) [36], sensor nodes share a common symmetric key with the base station that is kept hidden from intermediate aggregators. The major contribution of this work is the provision of end-to-end encryption for reverse multicast traffic between the sensors and the base station. In the proposed approach, data aggregators carry out aggregation functions that are applied to ciphertexts (encrypted data). This provides the advantage that intermediate aggregators do not have to carry out costly decryption and encryption operations. Therefore, data aggregators do not have to store sensitive cryptographic keys which ensures an unrestricted aggregator node election process for each epoch during the wireless sensor network's lifetime. Unrestricted data aggregator selection is impossible for hop-by-hop encryption because only the nodes which have stored the key can act as a data aggregator. As the privacy homomorphic encryption function, the proposed protocol employs the function proposed by Domingo-Ferrer [42]. Domingo-Ferrer's encryption function is probabilistic in the sense that the encryption transformation involves some randomness that chooses the ciphertext corresponding to a given plaintext from a set of possible ciphertexts.

The public parameters of Domingo-Ferrer's encryption function are a positive integer $d \geqslant 2$ and a large integer $g$ that must have many small divisors. In addition, there should be many integers less than g that can be inverted modulo $g$. The secret key is computed as $k = (r, g')$. The value $r \in \mathbb{Z}_g$ is chosen such that $r^{-1}$ mod $g$ exists where $log_{g'}g$ indicates the security level provided by the function. The set of plaintext is $\mathbb{Z}_{g'}$ and the set of ciphertext is $(\mathbb{Z}_g)^d$. Encryption and decryption processes are defined as follows:

*Encryption*: Randomly split $a \in \mathbb{Z}_{g'}$ into a secret $a_1 \cdots a_d$ such that $\sum_{j=1}^{d}(a_j mod g')$ and $a \in \mathbb{Z}_{g'}$. Compute

$$E_k(a) = (a_1 r^1 mod g, a_2 r^2 mod g, \cdots, a_d r^d mod g)$$

*Decryption:* Compute the $j$th coordinate by $r^{-j}$ mod $g$ to retrieve $a_j$ mod g. In order to obtain $a$ compute

$$D_k(E_k(a)) = \sum_{j=1}^{d}(a_j mod g')$$

The ciphertext operation $\times$ is performed by cross-multiplying all terms in $Z_g$, with the $d_1$-degree term by a $d_2$-degree term yielding a $t$-degree term. Then, the terms having the same degree are added up. The ciphertext operation + is relatively easy compared to $\times$ operation and is performed component-wise.

As it is seen from the above definitions, Domingo-Ferrers asymmetric key based privacy homomorphism is computationally expensive for resource constrained sensor nodes. Authors of [36] compared the clock cycles required by asymmetric key based privacy homomorphism and symmetric key based encryption solutions. The results show that encryption, decryption, and addition operations that are needed to implement Domingo-Ferrers function are much more expensive compared to those are necessary to perform symmetric key based RC5. However, the authors argue that this disadvantage is acceptable as CDA advantageously balance the energy consumption. Using symmetric key based encryption solutions to perform hop-by-hop data aggregation results in shorter lifetime for data aggregator nodes. Therefore, as data aggregators are the performance bottleneck when maintaining a connected wireless sensor network backbone, it is preferable to employ CDA's asymmetric key based privacy homomorphism to balance the energy consumption of data aggregators.

In [40], a secure data aggregation protocol, called CDAP, takes advantage of asymmetric key based privacy homomorphic cryptography to achieve end-to-end data confidentiality and data aggregation together. The authors point out that asymmetric cryptography based privacy homomorphism incurs high computational overhead which cannot be afforded by regular sensor nodes with scarce resources. To mitigate this problem, CDAP protocol employs a set of resource-rich sensor nodes, called *aggregator nodes* (AGGNODEs), for privacy homomorphic encryption and aggregation of the encrypted data. In CDAP, after the network deployment each AGGNODE establishes pairwise keys with its neighboring nodes so that neighboring nodes can send their sensor readings securely. In data collection phase of protocol CDAP, each AGGNODE queries its neighboring nodes. Each neighboring node encrypts its data (using RC5 algorithm) sends the encrypted data to its AGGNODE. The AGGNODE decrypts all the data received from its neighbors, aggregates them, and encrypts the aggregated data using the privacy homomorphic encryption algorithm. Once the data are encrypted with the privacy homomorphic encryption algorithm, only the base station can decrypt them using its private key. Due to homomorphic property, intermediate AGGNODEs can aggregate those encrypted data during data forwarding. Therefore, the data collected by sensor nodes are aggregated by AGGNODEs as they travel towards the base station. The base station decrypts the final aggregated data using its private key. An illustrative example of data aggregation in CDAP is given in Fig. 9. Due to the computational overhead of privacy homomorphic encryption algorithms, in CDAP, only AGGNODEs are allowed to encrypt and aggregate the collected data using privacy homomorphic algorithms. Therefore, during the initial data collection phase of the protocol CDAP, sensor nodes uses symmetric key algorithms for encryption. Due to the symmetric encryption, a compromised AGGNODE may disclose the secrecy of its neighboring nodes' data or inject false data into the data. However, the authors argue that the effect of this attack is local, and hence, it can be tolerated.

In [37], a simple and provably secure additively homomorphic stream cipher that allows efficient aggregation of encrypted data is proposed. The proposed technique is based on an extension of the one-time pad encryption technique using additive operations over modulo $n$. The main idea of the proposed scheme is to replace the $Exclusive - OR$ operation of stream ciphers with modular addition $(+)$. The encryption and decryption processes can be summarized as follows. Represent message $m$ as integer $m \in [0, M-1]$ where $M$ is a large integer. Also, let $k$ be a randomly generated key stream, where $k \in [0, M-1]$. Then, chiphertext $c$ is computed as $c = enc(m, k, M) = m + k(modM)$. In order to decrypt ciphertext $c$, perform $Dec(c, k, M) = c - k(modM)$. Based on these functions, ciphertexts are added as follows: Let $c_1 = Enc(m_1, k_1 M)$ and $c_2 = Enc(m_2, k_2, M)$, then for $k = k_1 + k_2, Dec(c_1 + c_2, k, M) = m_1 + m_2$. It is assumed that the message $m$ is $0 \leqslant m < M$ and since addition posses the commutative property, the proposed scheme is additively homomorphic. The proposed scheme significantly reduces the energy consumption of sensor nodes due to encryption process. However, in the proposed scheme, each aggregate message is coupled with the list of nodes that failed to contribute to the aggregation. When the aggregation tree is large, the list of sensor nodes become larger and results in a significant communication overhead. This problem has been solved

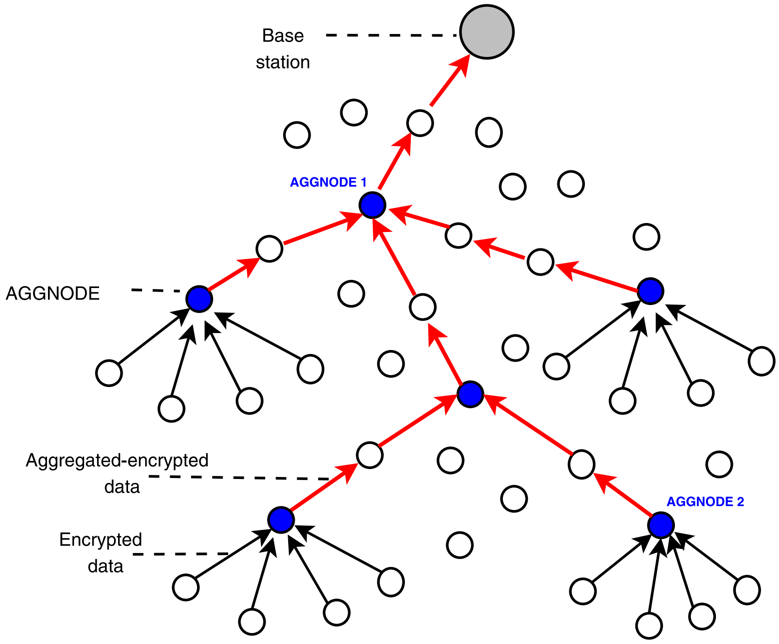

**Fig. 9.** The aggregation scenario of protocol CDAP. AGGNODEs collect information from their neighborhood and encrypted data are aggregated at AGGNODEs while data travels towards the base station.

in [38] by adapting a hierarchical data aggregation model. Similar to [37,38], a layered secure data aggregation protocol in wireless sensor networks that offers end-to-end data confidentiality by using homomorphic functions and interleaved encryption is proposed in [39]. The proposed protocol ensures that, in the presence of less than $n$ compromise nodes, an attacker cannot get access to any aggregated data from the network. When more than $n$ nodes are captured, the attacker can only get access to the aggregated values received by the captured nodes.

In [41], the authors realize the fact that existing privacy homomorphism based in-network processing protocols can only work for some specific query-based aggregation functions, e.g., sum, average, etc. Hence, instead of privacy homomorphism, the authors take advantage of digital watermarking and propose an end-to-end, statistical approach for data authentication that provides inherent support for data aggregation. The novel idea of this work is to modulate authentication information as watermark and superpose this information on the sensory data at the sensor nodes. The watermarked data can be aggregated by the intermediate nodes without incurring any en route checking. In order to check whether the data has been altered by the compromised nodes, upon reception of the sensory data, the data sink is able to authenticate the data by validating the watermark. More specifically, the proposed technique visualizes the sensory data gathered from the whole network at a certain time snapshot as an image, in which every sensor node is viewed as a pixel with its sensory reading representing the pixels intensity. Since senor data is represented as an "image" digital watermarking can be applied to this image. In order to balance the energy consumption among sensor nodes, a direct spread spectrum sequence (DSSS) based watermarking technique is used. While each sensor node appends a part of the whole watermark into its sensory data, verification of watermark which requires an extensive computational resource is left to the sink. The proposed scheme adopts the existing image compression schemes as the aggregation functions to reduce network load while retaining the desired details of the data. Moreover, using a DSSS based watermarking scheme, the proposed technique is enabled to survive a certain degree of distortion and therefore naturally support data aggregation.

Table 1 presents the comparison of secure data aggregation schemes with respect to wireless sensor network security requirements. As seen from Table 1, almost all secure data aggregation protocols ensure data integrity and source authentication. Protocols in [36,40,37,39] focus solely on aggregation of encrypted data and do not provide data integrity and source authentication support. However, these protocols can be modified easily to support data integrity and source authentication. Table 1 also shows that some of the secure data aggregation protocols ([26,30,31,34,35,41]) do not support data confidentiality which is essential for mission critical wireless sensor network applications. Therefore, these protocols should be used only in applications in which the transmitted data is not secret. Among the protocols that provide data confidentiality, the protocols proposed in [29,32,36,37,40,49] can offer end-to-end data confidentiality. It is also seen from Table 1 that only secure data aggregation protocols [34,35] support the availability of data aggregators. In order to achieve node availability protocols in [34,35] employ extensive monitoring mechanisms. Considering that data confidentiality, data integrity, and source authentication are the most important security requirements, we can conclude from Table 1 that data aggregation protocols proposed in [27–29,32,33,49] provide better security compared to other data aggregation protocols.

**Table 1**
Comparison of secure data aggregation protocols.

| Protocol | Data confidentiality | Data integrity | Source authentication | Node availability |
|---|---|---|---|---|
| Hu et al. [26] | | √ | √ | |
| SIA [27] | √ | √ | √ | |
| SecureDAV [28] | √ | √ | √ | |
| ESPDA [29] | √ | √ | √ | |
| Du et al. [30] | | √ | √ | |
| Wu et al. [31] | | √ | √ | |
| SRDA [32] | √ | √ | √ | |
| SDAP [33] | √ | √ | √ | |
| SELDA [34] | | √ | √ | √ |
| Ozdemir [35] | | √ | √ | √ |
| CDA [36] | √ | | | |
| Castelucia et al. [37] | √ | | | |
| Ozdemir [40] | √ | | | |
| Zhang et al. [41] | | √ | | |
| Rodhea et al. [39] | √ | | | |
| Sun et al. [47,48] | | √ | √ | |
| DAA [49] | √ | √ | √ | |

## 5. Open research issues and future research directions

In this paper, we present a comprehensive overview of secure data aggregation concept in wireless sensor networks. We survey the state-of-the-art data aggregation protocols and categorized them based on network topology and security. Although the presented research addresses the many problems of data aggregation, there are still many research areas that needs to be associated with the data aggregation process, especially from the security point of view.

As for the general data aggregation concept, the relation between routing mechanisms and data aggregation protocols have been well studied as they are highly correlated topics. In addition to diffusion and tree-based data aggregation protocols, many cluster-based data aggregation protocols that route aggregated data over cluster heads have been proposed. Although, these protocols shown to be very efficient in static networks in which the cluster structures do not change for a sufficiently long time, in dynamic networks they perform quite poorly. Hence, data aggregation in dynamic environments is a possible future research direction. The impact of sensor node heterogeneity over the data aggregation protocols is another unexplored research area [40]. The protocols that use powerful sensor nodes as data aggregators presented promising results. However, determining locations of these powerful nodes for the best data aggregation results needs further research.

Security is an important issue for data aggregation process and it needs to be further investigated. Clearly, there are still secure data aggregation issues that have not been addressed by the existing research. One such problem is compromised data aggregators that inject false data during data aggregation. Because data aggregation usually results in alterations in collected sensor data, false data injections by compromised data aggregators are hard to detect. There is only limited work targeting this problem and the proposed techniques are all based on extensive node monitoring mechanisms [47–49]. The efficiency of these node monitoring protocols is not fully evaluated and they usually incur high radio and sensing resource consumption. Hence, development of lightweight monitoring mechanisms specifically for secure data aggregation process is an interesting problem for future research.

In order to provide end-to-end security, privacy homomorphism based secure data aggregation protocols have drawn considerable attention recently. However, the design and implementation of resource efficient privacy homomorphic aggregation functions yet to be explored. Many existing public key cryptography based privacy homomorphic functions are not feasible for resource limited sensor nodes. Hence, in some secure data aggregation schemes elliptic curve cryptography is employed [36]. However, these elliptic curve cryptography based privacy homomorphic functions can only work for some specific query-based aggregation functions, e.g., sum, average, etc. Therefore, design of efficient privacy homomorphic functions that are able to work with all types of data aggregation functions needs to be explored. In addition, for certain wireless sensor network settings where real-time data delivery is demanded, symmetric key cryptography based privacy homomorphic encryption schemes are recommended [38,37]. But, there are not many symmetric key based privacy homomorphic schemes. Hence, exploration of symmetric key cryptography based privacy homomorphic functions in the secure data aggregation concept is another promising research area. Using “digital watermarking” schemes to replace the expensive privacy homomorphic functions is a newly introduced concept in secure data aggregation [41]. However, this method allows only one way authentication of sensor data at the base station. Hence, investigation of two-way authentication by using watermarking techniques that will allow in-network secure data aggregation in the network may be a good research direction.

In addition, the application of source coding theory for data aggregation has drawn a little attention so far. Considering that sensor data is highly correlated, data aggregation can be achieved by employing source coding techniques. Existing research in this area focuses on only theoretical results and there are no practical algorithms applicable to wireless sensor networks yet. Moreover, there is no secure data aggregation protocol that uses the idea of source coding which may seamlessly integrate data confidentiality and aggregation together. Therefore, there is significant scope for future work in source coding based secure data aggregation.

Secure hierarchical data aggregation is expected to produce a vast amount of research in the future. Many secure data aggregation protocols assume that sensor data are aggregated at a single sink or data aggregator. Especially for privacy homomorphic secure data aggregation protocols providing hierarchical aggregation is not a trivial task. Hence, extending the current single level secure data aggregation protocols to multi layer hierarchical data aggregation protocols is an interesting problem for future research.

## 6. Conclusion

This paper provides a detailed review of secure data aggregation concept in wireless sensor networks. To give the motivation behind secure data aggregation, first, the security requirements of wireless sensor networks are presented and the relationships between data aggregation concept and these security requirements are explained. Second, an extensive literature survey is presented by summarizing the state-of-the-art data aggregation protocols. Based on this extensive literature survey, open research areas and future research directions are given.

## References

[1] I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, A survey on sensor networks, IEEE Commun. Mag. 40 (8) (2002) 102–114.

[2] J. Yick, B. Mukherjee, D. Ghosal, Wireless sensor network survey, Comput. Networks 52 (12) (2008) 2292–2330.

[3] K. Akkaya, M. Demirbas, R.S. Aygun, The Impact of Data Aggregation on the Performance of Wireless Sensor Networks, Wiley Wireless Commun. Mobile Comput. (WCMC) J. 8 (2008) 171–193.

[4] J. Newsome, E. Shi, D. Song, A. Perrig, The Sybil attack in sensor networks: analysis and defenses, in: Proceedings of the Third IEEE/ACM Information Processing in Sensor Networks (IPSN'04), 2004, pp. 259–268.

[5] A. Perrig, R. Szewczyk, D. Tygar, V. Wen, D. Culler, SPINS: security protocols for sensor networks, Wireless Networks J. (WINE) 2 (5) (2002) 521–534.

[6] Crossbow Technologies Inc. <http://www.xbow.com>.

[7] E. Fasolo, M. Rossi, J. Widmer, M. Zorzi, In-network aggregation techniques for wireless sensor networks: a survey, IEEE Wireless Commun. 14 (2) (2007) 70–87.

[8] R. Rajagopalan, P.K. Varshney, Data aggregation techniques in sensor networks: a survey, IEEE Commun. Surveys Tutorials 8 (4) (2006).

[9] C. Intanagonwiwat, D. Estrin, R. Govindan, J. Heidemann, Impact of network density on data aggregation in wireless sensor networks, in: Proceedings of the 22nd International Conference on Distributed Computing Systems, 2002, pp. 457–458.

[10] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, F. Silva, Directed diffusion for wireless sensor networking, in: IEEE/ACM Transactions on Networking, vol. 11, 2003, pp. 2–16.

[11] B. Krishnamachari, D. Estrin, S. Wicker, The impact of data aggregation in wireless sensor networks, in: Proceedings of the 22nd International Conference on Distributed Computing Systems Workshops, 2002, pp. 575–578.

[12] M. Ding, X. Cheng, G. Xue, Aggregation tree construction in sensor networks, in: Proceedings of the 58th IEEE Vehicular Technology Conference, vol. 4, 2003, pp. 2168–2172.

[13] R. Cristescu, B. Beferull-Lozano, M. Vetterli, On network correlated data gathering, in: Proceedings of the 23rd Annual Joint Conference of the IEEE Computer and Communications Societies, vol. 4, 2004, pp. 2571–2582.

[14] S. Madden et al., TAG: A Tiny AGgregation Service for Ad Hoc Sensor Networks, OSDI, Boston, MA, 2002.

[15] B. Zhou et al., A Hierarchical Scheme for Data Aggregation in Sensor Network, IEEE ICON 04, Singapore, 2004.

[16] M. Lee, V.W.S. Wong, An Energy-Aware Spanning Tree Algorithm for Data AggrAegation in Wireless Sensor Networks, IEEE PacRrim, Victoria, BC, Canada, 2005.

[17] S. Lindsey, C. Raghavendra, K.M. Sivalingam, Data gathering algorithms in sensor networks using energy metrics, IEEE Trans. Parallel Distrib. Sys. 13 (9) (2002) 924–935.

[18] G. Di Bacco, T. Melodia, F. Cuomo, A MAC Protocol for Delay-Bounded Applications in Wireless Sensor Networks, Med-Hoc-Net, Bodrum, Turkey, 2004.

[19] W.B. Heinzelman, A.P. Chandrakasan, H. Balakrishnan, An application-specific protocol architecture for wireless micro-sensor networks, IEEE Trans. Wireless Commun. 1 (4) (2002) 660–670.

[20] O. Younis, S. Fahmy, HEED: a hybrid, energy-efficient distributed clustering approach for ad hoc sensor networks, IEEE Trans. Mobile Comput. 3 (4) (2004) 366–379.

[21] Y. Yao, J. Gehrke, The Cougar approach to in-network query processing in sensor networks, ACM SIGMOD Rec. 31 (3) (2002) 9–18.

[22] S. Chatterjea, P. Havinga, A dynamic data aggregation scheme for wireless sensor networks, in: Proceedings of the Program for Research on Integrated Systems and Circuits, Veldhoven, The Netherlands, 2003.

[23] V. Mhatre, C. Rosenberg, Design guidelines for wireless sensor networks: communication clustering and aggregation, Elsevier Ad Hoc Networks J. 2 (1) (2004) 45–63.

[24] P. Popovski et al., MAC-Layer Approach for Cluster-Based Aggregation in Sensor Networks, IEEE IWWAN 04, Oulu, Finland, 2004.

[25] S. Pattem, B. Krishnamachari, R. Govindan, The Impact of Spatial Correlation on Routing with Compression in Wireless Sensor Networks, ACM/IEEE IPSN04, Berkeley, CA, 2004.

[26] L. Hu, D. Evans, Secure aggregation for wireless networks, in: Proceedings of the Workshop on Security and Assurance in Ad Hoc Networks, Orlando, FL, 28 January 2003.

[27] B. Przydatek, D. Song, A. Perrig, SIA : secure information aggregation in sensor networks, in: Proceedings of SenSys'03, 2003, pp. 255–265.

[28] A. Mahimkar, T.S. Rappaport, SecureDAV: a secure data aggregation and verification protocol for wireless sensor networks, in: Proceedings of the 47th IEEE Global Telecommunications Conference (Globecom), November 29–December 3, Dallas, TX, 2004.

[29] H. Çam, S. Ozdemir, P. Nair, D. Muthuavinashiappan, H.O. Sanli, Energy-efficient and secure pattern based data aggregation for wireless sensor networks, Comput. Commun., Elsevier 29 (4) (2006) 446–455.

[30] W. Du, J. Deng, Y.S. Han, P.K. Varshney, A witness-based approach for data fusion assurance in wireless sensor networks, in: Proceedings of the IEEE Global Telecommunications Conference (GLOBECOM '03), 2003, pp. 1435–1439.

[31] K. Wu, D. Dreef, B. Sun, Y. Xiao, Secure data aggregation without persistent cryptographic operations in wireless sensor networks, Ad Hoc Networks 5 (1) (2007) 100–111.

[32] H.O. Sanli, S. Ozdemir, H. Çam, SRDA: secure reference-based data aggregation protocol for wireless sensor networks, in: Proceedings of the IEEE VTC Fall Conference, Los Angeles, CA, 26–29 September 2004, pp. 4650–4654.

[33] Y. Yang, X. Wang, S. Zhu, G. Cao, SDAP: a secure hop-by-hop data aggregation protocol for sensor networks, in: Proceedings of the ACM MOBIHOC'06, 2006.

[34] S. Ozdemir, Secure and reliable data aggregation for wireless sensor networks, in: H. Ichikawa et al. (Eds.), LNCS 4836, 2007, pp. 102–109.

[35] S. Ozdemir, Functional reputation based reliable data aggregation and transmission for wireless sensor networks, Elsevier Comput. Commun. 31 (17) (2005) 3941–3953.

[36] D. Westhoff, J. Girao, M. Acharya, Concealed data aggregation for reverse multicast traffic in sensor networks: encryption key distribution and routing adaptation, IEEE Trans. Mobile Comput. 5 (10) (2006) 1417–1431.

[37] C. Castelluccia, E. Mykletun, G. Tsudik, Efficient aggregation of encrypted data in wireless sensor networks, in: Proceedings of the Conference on Mobile and Ubiquitous Systems: Networking and Services, 2005, pp. 109–117.

[38] S. Ozdemir, Secure data aggregation in wireless sensor networks via homomorphic encryption, Journal of The Faculty of Engineering and Architecture of Gazi University 23 (2) (2008) 365–373. ISSN:1304-4915.

[39] I. Rodhea, C. Rohner, n-LDA: n-layers data aggregation in sensor networks, in: Proceedings of the 28th International Conference on Distributed Computing Systems Workshops, 2008, pp. 400–405.

[40] S. Ozdemir, Concealed data aggregation in heterogeneous sensor networks using privacy homomorphism, in: Proceedings of the ICPS'07: IEEE International Conference on Pervasive Services, Istanbul, Turkey, 2007, pp. 165–168.

[41] W. Zhang, Y. Liu, S.K. Das, P. De, Secure data aggregation in wireless sensor networks: a watermark based authentication supportive approach, Elsevier Pervasive Mobile Comput. 4 (2008) 658–680.

[42] J. Domingo-Ferrer, A provably secure additive and multiplicative privacy homomorphism, in: Proceedings of the Information Security Conference, 2002, pp. 471–483.

[43] T. Okamoto, S. Uchiyama, A New Public-Key Cryptosystem as Secure as Factoring, Advances in Cryptology – EUROCRYPT'98, 1998, pp. 208–318.

[44] A. Josang, R. Ismail, The beta reputation system, in: Proceedings of the 15th Bled Conference Electronic Commerce, 2002.

[45] S. Ganeriwal, M.B. Srivastava, Reputation-based framework for high integrity sensor networks, in: Proceedings of the Second ACM Workshop on Security of Ad Hoc and Sensor Networks, Washington DC, 2004, pp. 66–77.

[46] Y. Xu, J. Heidemann, D. Estrin, Geography-informed energy conservation for ad hoc routing, in: Proceedings of the CM/SIGMOBILE MobiCom, 2001, pp.70–84.

[47] B. Sun, X. Jin, K. Wu, Y. Xiao, Integration of secure in-network aggregation and system monitoring for wireless sensor networks, in: Proceedings of IEEE International Conference on Communications (IEEE ICC'07), 2007, pp. 1466–1471.

[48] B. Sun, N. Chand, K. Wu, Y. Xiao, Change-point monitoring for secure in-network aggregation in wireless sensor networks, in: Proceedings of IEEE Global Telecommunications Conference, IEEE GLOBECOM, 2007, pp. 936–940.

[49] H. Çam, S. Ozdemir, False data detection and secure data aggregation in wireless sensor networks, in: Yang Xiao (Ed.), Security in Distributed Grid Mobile and Pervasive Computing, Auerbach Publications, CRC Press, 2007.

[50] B. Parekh, H. Çam, Minimizing false alarms on intrusion detection for wireless sensor networks in realistic environments, in: Proceedings of IEEE Military Communications Conference, 2007, pp. 1–7.

**Suat Ozdemir** has been with the Computer Engineering Department at Gazi University, Ankara, Turkey since March 2007. He received his M.Sc. degree in Computer Science from Syracuse University (August 2001) and Ph.D. degree in Computer Science from Arizona State University (December 2006). His main research interests include broad areas of wireless networks and network security. He serves on TPC for several conferences such as ICC, GLOBECOM, Sensor Networks, Information Security and Cryptography, etc. He also serves as a reviewer for several journals, e.g., IEEE Transactions on Mobile Computing, IEEE Transactions on Parallel and Distributed Systems, Computer Communications, Computer Networks.

**Yang Xiao** worked in industry as a MAC (Medium Access Control) architect involving the IEEE 802.11 standard enhancement work before he joined Department of Computer Science at The University of Memphis in 2002. He is currently with Department of Computer Science at The University of Alabama. He was a voting member of IEEE 802.11 Working Group from 2001 to 2004. He is an IEEE Senior Member. He is a member of American Telemedicine Association. He currently serves as Editor-in-Chief for International Journal of Security and Networks (IJSN), International Journal of Sensor Networks (IJSNet), and International Journal of Telemedicine and Applications (IJTA). He serves as a referee/reviewer for many funding agencies, as well as a panelist for NSF and a member of Canada Foundation for Innovation (CFI)'s Telecommunications expert committee. He serves on TPC for more than 100 conferences such as INFOCOM, ICDCS, MOBIHOC, ICC, GLOBECOM, WCNC, etc. He serves as an associate editor for several journals, e.g., IEEE Transactions on Vehicular Technology. His research areas are security, telemedicine, sensor networks, and wireless networks. He has published more than 300 papers in major journals, refereed conference proceedings, book chapters related to these research areas. His research has been supported by the US National Science Foundation (NSF), US Army Research, Fleet and Industrial Supply Center San Diego (FISCSD), and The University of Alabama's Research Grants Committee.