An Investigation Of Secure And Energy Efficient Data Aggregation In Wireless Sensor Networks

Zehra Karapinar Senturk, Arafat Senturk, Resul Kara

Abstract: The most energy consuming operation in a wireless sensor network is data transmission. The bigger data to be transmitted, the more energy is consumed. Therefore, minimizing the amount of data to be transferred is very important. Data aggregation appears at this point. Summary of the data of a group of node is transmitted to sink or another node instead of sending all data. Since data aggregation is that much important, it is usual to see some counter attacks to overshadow the network process. Also, it is natural to see some studies to increase the efficiency of data aggregation process in the literature. In this paper, we give an overview about recent secure and energy efficient data aggregation studies. We also referred attack types faced in wireless sensor networks and give some recommendations to cope with them.

Index Terms: Data aggregation, security, attacks, energy efficiency.

1 INTRODUCTION

Wireless sensor networks are becoming more and more prevalent nowadays because of their convenience in establishment and usage. It is very usual to see them in almost every area around us from healthcare monitoring [1] to object tracking [2]. Sensors used in the network are inexpensive and small devices, but they have very limited battery, communication capability, and memory space. With these drawbacks, wireless sensor networks are prone to several malign attacks like denial of service attacks, false data injection attacks, etc. Therefore, it is vitally important for a wireless sensor network to be robust against those and other attacks. One of the most important processes in a sensor network in terms of energy consumption and security is data aggregation. Data aggregation is just summarizing data before sending it to a sink or another node to decrease the energy consumed during data transmission, and so, increase the lifetime of the network. Finite energy sources of the sensors entail such kind of optimization efforts to get best result from the network. For the security of data aggregation, both trustworthiness of aggregator nodes and the trustworthiness of non-aggregator nodes must be considered. Because, for example, non-aggregator nodes may send true sensing data to their aggregator and the aggregator may change it before relaying it to an upper level or aggregator can be reliable but the nodes which send data to it may send malignant data.

- *Zehra KARAPINAR SENTURK, PhD student in Sakarya University, Turkey, E-mail: zehrakarapinar@duzce.edu.tr*
- *Arafat SENTURK, PhD student in Duzce University, Turkey, E-mail: arafatsenturk@duzce.edu.tr*
- *Resul KARA, Assoc. Prof. in Duzce University, Turkey, E-mail: resulkara@duzce.edu.tr*

Outsider attacks must also be considered in terms of data aggregation besides those mentioned insider attacks to provide a fully secure data aggregation. This paper is organized as follows: In section 2, we mentioned about the studies which focus on secure and energy efficient data aggregation in the literature. Section 3 is composed of attack types which threaten the network safety and the precautions against those. Section 4 concludes the paper and gives some recommendations and future work.

2 RELATED WORKED

Providing an appropriate security level is a burdensome complication in wireless sensor networks because they are resource and bandwidth limited [3]. Authentication, integrity of the messages, and the confidentiality of data must be provided for a secure network. There are many studies on these security mechanisms. [3] highlights the fact that trust between two communicating nodes is critical and gives a summary of some node and data trust models. Six node trust models (TCFL, RFSN, PLUS, NBBTR, ATSN, and TTSN) and four data trust models (DFDI, DFR, MDLC, and TMCDE) are evaluated. Every mechanism has some limitations like requiring centralized management, not to be able to improve system robustness but improving individual node security, high trust convergence time, high energy, time, and memory costs, vulnerability to malicious agent nodes, unrealistic trust values, and vulnerability to collusion attacks. Recently, many mobile agent based scheme uses static itineraries which are computed at BS using a centralized algorithm for agent's migration and these algorithms require global knowledge of sensor distribution which may not always be known. Another problem is that agents may not move along their itineraries due to node failures or malicious node attacks. A framework for trust evaluation to identify malicious node behavior and a localized distributed protocol, called energy and trust aware mobile agent migration (ETMAM) protocol for periodic data gathering application are proposed in [4]. The efficiency and effectiveness of mobile agent based data aggregation depend on the agent's migration path either dynamic or static. Static approach brings significant additional cost since it needs periodic collection of network topology information at BS and an agent may not move along its path due to node failures or malicious attacks. Dynamic approach is more flexible in case of node or link failures and its agent packet size is smaller than the static one since it doesn't carry pre-computed itineraries list. Protection of mobile agents against malicious nodes is another critical issue besides energy efficiency. Basic

cryptography methods like symmetric key authentication and integrity protection are not enough. Because compromised nodes know secret keys and may behave as if they are loyal. Public-key cryptography involves considerable storage and computation overheads and not applicable to WSNs. So, to give an efficient solution to overcome these attacks, trustworthiness of nodes should be used to identify and bypass malicious nodes during agent's migration along the network. The protocol proposed in [4] combines energy and trust to establish routes for travelling agent to complete data aggregation tasks. Trust evaluation is realized by both peer recommendation and local monitoring of node's behavior. ETMAM uses agent cloning concepts and decides next-hop for agent migration using energy and trust value of next-hop node. If trust value of next-hop node is below threshold, then it is not selected as next-hop node and route is modified. So, malicious nodes are skipped. ETMAM facilitates small itinerary length for each agent thanks to agent cloning features and has lower communication cost. Since ETMAM's agents work in parallel, its response time is low. The weights of the nodes inside the network are thought to be different in [5] and they contribute the aggregation proportional to their weights. An ID-based secure lossy data aggregation integrity scheme is proposed. Homomorphic hashing and ID-based aggregate signature is used to provide security. All aggregators beside the sink node are able to verify the authenticity of aggregated data via a distinct key which is shared by the sink and every node. Aggregator nodes and BS are aware of the weights of sensor nodes. Data integrity in data aggregation is intended to be supplied in a cryptographic manner and some problems are aimed to be solved such that need of base station to know weights of non-uniform sensor nodes, key security, and the case where distinct private keys are used in multiple sensor nodes. Solution to that problem is intended to be solved by public-key cryptography and ID-based data integrity. Firstly, hash values of sensor readings are computed and those hash values are signed. Then, hash values and signatures are sent together with the data to be sent to the aggregator. Lastly, base station authenticates the integrity without knowledge of private keys of sensors [5]. This approach is a generic one since it can be easily used when there are multiple private keys instead of a common key. Because the base station realizes integrity checking being unaware of the private keys of nodes. Also, each sensed data is divided into n blocks of equal lengths such and base station uses this sensed data vector to calculate the weights of sensor nodes. It is the linear combination of that vector. Base station needs nothing else for weight calculation. Trust level evaluation is realized by a central collector node, i.e. a cluster head, a gateway, etc. according to the scenario in [6]. Cluster head takes and evaluates the messages of loyal nodes while dropping the messages of betrayers. Central node collects data and verifies data authenticity, integrity, confidentiality, and freshness. Trust level calculation is done by Momani's model combined with Byzantine protocol. In the scenarios given in paper, even if two nodes cheat and verify each other (cooperative disloyalty), majority of the nodes in that cluster exposes the disloyalty of malicious nodes since their sensor readings are different (above/below the threshold value) than others. When the environment changes, i.e. sensors read different values at a consecutive time slice, system accepts all nodes to be loyal at first. In the first trust value calculation, readings of sensors will not be verified. So, their trust values become dropping. But after recalculation phase, system adopts itself to new condition in about 3000ms. A protocol, called DAA, is proposed in [7] to detect false data aggregation and to provide confidentiality. Monitoring nodes are employed to realize aggregation, too. They compute MAC and send to their pair mates. So, data verification is provided. Pair wise and group keys are assumed to be established. Pairwise keys are used between monitoring nodes and their pair mates. Sybil attacks are prevented by ensuring the IDs of pair mates of monitoring nodes. Group keys are used for two purposes. First, it is used to select monitoring nodes. They are also used to protect data confidentiality between aggregator and neighbors for data verification and aggregation. The selection of monitoring nodes is realized by considering both aggregator and all neighboring nodes. The purpose of this is to minimize possibility of selecting a malicious node as aggregator. Detecting false data in aggregators rather than base station decreases the amount of data transmission and so, increase lifetime of the network. The protocol given in [8] provides confidentiality via homomorphic additive function. Each aggregator aggregates encrypted but not the original data sent by sensor nodes and send the encrypted aggregated result to the next aggregator or sink with a pairwise shared key. Each receiver extracts their pairwise shared key from gathered data and obtains just ciphered data and a random number. This process goes on until arriving to base station. Base station takes ciphered data and decrypts it with its pairwise key. So, the encrypted data goes from leaf sensor nodes to base station without decrypting and decrypted only by base station which is the ideal aggregation process. Proposed protocol is compared by traditional data aggregation protocols and by a protocol which does not make aggregation to send data to sink. Results show that this new method decreases energy consumption of the network and provides end-to-end data confidentiality. [10] proposes secure information aggregation protocol for WSNs and uses only one aggregator which is the sink. In this protocol, base station collects all authenticated data and aggregates them and then, sends the aggregated data to a secure end user with a commitment. There exists one main aggregator and multiple witness aggregators in [11]. Reliability of the aggregator is checked by its witnesses and so, the integrity of aggregation result is provided. To provide a secure communication in a wireless sensor network, nodes in [12] uses private and pair wise shared keys. Private keys are used by aggregators to sign aggregation results not to allow alterations and pair wise shared keys are used by the nodes in each link in addition to private keys. It is assumed that ID-based public key crypto-system exists. Every node has a private key and other nodes are able to calculate that node's public key with its ID. Data confidentiality and data integrity attacks in time sensitive WSNs are tried to be defended in [13] with a low-cost data communication. RF communication channel is benefitted in this approach. Passive participation is supported and nodes are able to hear the submissions of their neighbors. Therefore, base station can detect data drops thanks to adversary's neighbor. Data confidentiality is provided via a secret key shared by sink and sensors. MAC is used for message integrity. A different privacy protection mechanism is given in [14]. Authors of that paper considered to protect the network against eavesdropping. Leaf nodes slice their readings and send that pieces to different neighbors randomly. Then, they receive slices and combine them before sending to aggregator, their parents. Aggregators also read data and they

aggregate their own readings and the data coming from child nodes. Therefore, they cannot conceal original sensor reading. A comparison is done with SMART (Slice-Mix-AggRegaTe) which also provides privacy via slicing and assembling technique in terms of bandwidth consumption and shown that their approach is better than SMART.

3 Attack Types and Countermeasures

Data aggregation is one of the most important parts of sensor networks since it seriously reduces the amount of data to be transmitted. It saves most of the energy of the network and provides a longer lifetime. Previously mentioned studies are developed to provide a secure and efficient network mostly by resisting against attacks. Attack types [3] and the countermeasures are given below.

DoS attack: the purpose in DoS attack is to use up energy and memory sources of the counterpart. That kind of an attack to an aggregator node is quite risky for a network. Because losing an aggregator node means losing several nodes which send their sending data directly to that aggregator. So, the communication with a region may be broken down. During the attack many ambiguous data is sent continuously to the target and the target body is forced to accept data frequently. Those ambiguous data fills the memory of the target unduly and its battery discharges soon. [3] proposes that this problem can be solved with the power-aware trust models. Another solution can be accepting certain amount of data in a certain time slice. If more than that amount of data is sent to aggregator, they are not regarded and the node which sends data frequently can be punished (the weight of its data may be decreased) or it can be marked as malicious and never regarded any more.

Bad mouthing attack: in this attack, deceitful nodes give incorrect recommendation willfully about their neighbors. Loyal neighbors are tried to be blackened by those malicious nodes. Propagating only good reputation information about other nodes is proposed as a solution to that attack in [3]. But this is not logical in case there is more than one malicious node in the network. Those malicious nodes may always give positive recommendations about their adherents. [3] also tells that trust models based on direct neighbor sensing or aggregations of multiple observations can handle well. In addition, aggregators may keep information tables about the nodes they communicate. If some nodes (less than the half of total node number) give bad information about a node in the network while most of the nodes give good recommendation, then those nodes which give bad recommendation can be punished and/or marked as malicious. This mechanism can also be used for preventing from **collusion attack**s when the number of malicious nodes is less than half of total node number in the network.

On-off attack: the behaviors of malicious nodes are inconsistent in this attack type. They are sometimes loyal and sometimes disloyal according to their benefits. This attack type is one of the hardest ones to be resisted because aggregator may remain trusted while attacker behaves abjectly. In the literature there are suggestions like using a forgetting factor in observations [15] and aggregations of multiple observations [16, 17, and 18]. An improvement to suggestions can be keeping only last data sent by a node and if that node sends very different data, then only at this time aggregator applies the observations of other nodes. Instead of always applying to multiple observations, applying only in situation changes make the aggregator save more energy and so, increase the lifetime of the network.

Conflicting behavior attack: malicious nodes in this attack give different recommendations to different nodes about a node. This creates an ambiguity about the trustworthiness of that node. Solutions developed for on-off attack can also be used for conflicting behavior attack. Additionally, information tables may be kept by the aggregators and each aggregator is hold to account for the nodes which are inside its neighborhood. Therefore, leaf nodes can only send information about other leafs which are connected to the same aggregator. Recommendation is one-way and through only from leaf nodes to their aggregator (one node). So, one cannot give good recommendation about a node to a master while giving bad recommendation to another master about that node.

Sybil attack: some feint IDs are established and some nodes in the network are impersonated in this attack. Attacker may change the recommendations and declare itself as a trusted node. ID identification or centralized trust methods in which BS can detect fake IDs can cope with this. Newly joined node may be forced to select an ID from an ID pool and that selected ID can then be marked as ‘given’. That is, nodes to be joined to the network should choose an ID which is not marked as ‘given’. Therefore, malicious node cannot create a fake ID and impersonate the nodes in the network. This mechanism can protect against **replication attack**s, too.

Replication attack: when a node is captured and its keys are dismissed, its replicas are created by the attacker and those replicas are sprinkled around the network. This attack type is very similar to Sybil attack and the same protection mechanisms can be used.

Attack on information: malicious node may change, copy or snitch information in this attack [3]. This is one of the mostly faced attacks and several encryptions and keying mechanisms such as [8] can handle it well.

Collusion attack: this is more dangerous than other attacks since there are collaboratively working malicious nodes. The scenarios which depend on direct observation of each node can resist [15, 18]. Also, the solution suggested above for bad mouthing attack can be used.

4 Conclusions and Future Work

There exists an overview for secure and energy efficient data aggregation which is a quite important phase in an effective network. Attack models faced in wireless sensor networks are also mentioned and some suggestions are given to resist against those attacks. Some recommendations and future works can be summarized as:

- Authenticity, integrity, and confidentiality must be provided for a secure network.
- Selecting random nodes to prove trustworthiness of an aggregator can be effective.
- Passive participation must be considered with its security and energy risks.

- Grouping or grading sensor readings may prevent attacks of outsiders. Leaf sensor nodes read data and send only a grade or group information of that data to their aggregator. Since this grouping of sensed data is only known by the nodes inside the network, an outsider can never have an idea about eavesdropped data.
- Accepting certain amount of data in certain time slice can prevent from DoS attacks.
- Tree based aggregator transmission must be chosen instead of direct communication with sink in terms of energy efficiency.
- It is not an adequate evaluation to calculate only direct or indirect trust.

Trust computation must be designed as task dependent.

ACKNOWLEDGEMENT

This work is supported by Scientific Research Project Council of Duzce University with project number 2010.03.04.050.

References

[1] Y.M.Huang, M.Y.Hsieh, H.C.Chao, S.H.Hung, and J.H. Park, “Pervasive, secure access to a hierarchical-based health care monitoring architecture in wireless heterogeneous sensor networks”, IEEE Journal on Selected Areas of Communications, vol 4, pp. 400-411, 2009.

[2] H.T. Kung and D. Vlah, “Efficient Location Tracking Using Sensor Networks”, presented at Proceedings of 2003 IEEE Wireless Communications and Networking Conference (WCNC), 2003, 2003 IEEE. Vol. 3. IEEE, 2003.

[3] H. Guangjie, J. Jiang, L. Shu, J. Niu, and H.C. Chao, “Management and applications of trust in Wireless Sensor Networks: A survey, ”Journal of Computer and System Sciences”, vol. 80, pp. 602-617, May 2014.

[4] G.P. Gupta, M. Misra, and K. Garg, “Energy and trust aware mobile agent migration protocol for data aggregation in wireless sensor networks”, Journal of Network and Computer Applications”, vol. 41, pp. 300-311, May 2014.

[5] S. Niu, C.Wang, Z. Yu, S. Cao, “Lossy data aggregation integrity scheme in wireless sensor networks”, Computers & Electrical Engineering, vol. 39, pp. 1726-1735, August 2013.

[6] B. Stelte and A. Matheus, “Secure Trust Reputation with Multi-Criteria Decision Making for Wireless Sensor Networks Data Aggregation”, presented at Sensors, 2011 IEEE, Limerick, Oct. 28-31 2011, Paper 920 – 923.

[7] S. Ozdemir and H. Cam, “Integration of False Data Detection With Data Aggregation and Confidential Transmission in Wireless Sensor Networks”, IEEE/ACM Transaction on Networking”, vol. 18, pp. 736, 749, June 2010.

[8] S. Ozdemir, “Secure data aggregation in wireless sensor networks via homomorphic encryption”, Journal of the Faculty of Engineering and Architecture of Gazi University, vol. 23, pp. 365-373, 2008.

[9] S. Ozdemir and Y. Xiao, “Secure data aggregation in wireless sensor networks: A comprehensive overview”, Computer Networks”, vol. 53, pp. 2022-2037, August 2009.

[10] B. Przdatek, D. Song, and A. Perring, “SIA: secure information aggregation in sensor networks”, Proceeding of the First International Conference on Embedded Networked Sensor Systems, November 2003.

[11] W. Du, J, Deng, Y. Han, and P. K. Varshney, “A witness-based approach for data fusion assurance in wireless sensor networks”, Proceeding of the IEEE Global Telekominications Conference, 2003.

[12] H. Li, K. Li, W. Qu, and I. Stojmenovic, “Secure and energy-efficient data aggregation with malicious aggregator identification in wireless sensor networks”, Future Generation Computer Systems, vol. 37, pp. 108-116, 2014.

[13] Y. Zhao, Y. Zhang, Z. Qin, and T. Znati, “A co-commitment based secure data collection scheme for tiered wireless sensor networks”, Journal of System Architecture, vol. 57, pp. 655-662, June 2011.

[14] H. Li, K. Lin, and K. Li, “Energy-efficient and high-accuracy secure data aggregation in wireless sensor networks”, Computer Communications, vol. 34, pp. 591-597, April 2011.

[15] H. Ceng, H. Wun, X. Zhou, and C.Gao, “Agent-based trust model in wireless sensor networks”, 8th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel / Distributed Computing, pp. 119-124, 2007.

[16] Z. Yao, D. Kim, and y. Doh, “PLUS: Parameterized and localized trust management scheme for sensor network security”, IEEE International Conference on Mobile ad-hoc and Sensor System, MASS, pp. 437-446, 2008.

[17] R. Feng, X. Xu, X. Zhou, and J. Wan, “A trust evaluation algorithm for wireless sensor networks based on node behaviours and D-S evidence theory”, Sensors, vol. 11, pp. 1345-1360, 2011

[18] H. Chen, “Task-based trust management for wireless sensor networks”, International Journal of Security and its Applications 3, vol. 2, pp. 21-26, 2009.