# Security Implications of Data Dissemination Methods in Wireless Sensor Networks

Jeremy Langston
Electrical and Computer Engineering
Tennessee Technological University
Cookeville, TN 38505, USA
Email: jwlangston21@tntech.edu

*Abstract*

*The latest movement in sensory data collection uses wireless communications. Usage of wireless communications lends itself to lower costs in installations (e.g. labor, copper/fiber media) and allows the sensor locations to be more easily relocated. However, using an open medium in wireless data relay raises security issues. Since each sensor node is typically resource limited, the transmission range is also limited and, thus, data must hop from one node to the other. This paper looks at the problems with the dissemination of data between sensor nodes from the security aspects such as confidentiality, integrity, and availability. From the three types of dissemination methods, external, local, and data-centric, each are compared to each other using a set of security threats. Implications derived from this study paint a varying view as to which method is the most secure. While many factors are at play, external storage rises to the top due to its inherent characteristics, such as fundamentally not storing its data within the network.*

## I. Introduction

The field of wireless Sensor Networks (WSNs) has become a hot topic for research due to their many applications and challenges. A WSN is a collaboration of sensor nodes. These nodes sense their surroundings and take readings and measurements. They also build a communications network to send data that they sense. Sensor data ranges from temperature to humidity to proximity sensing to event captures. An example event may be compounding temperature and movement sensor data to determine the presence of a human.

Node deployment can expand endlessly. Previously, sensor networks were limited by two main obstacles: communication and power/energy limitations. For a typical sensor network, all nodes must tie back to the network using a land-line, or physical connection, for communication, as well as a power supply. WSNs circumvent these requirements by using wireless communications and batteries. Not only does this allow for decreased deployment costs, but it also allows for much greater flexibility. As long as the nodes have remaining energy and can communicate with a nearby node, a node can be placed anywhere.

Typical applications include environmental concerns and military deployment, but are not limited to these. Consider the summer months in southern California. Flash fires can cause millions of dollars of damage to residential and commercial buildings. With a wide-area sensor net can be preemptively arranged in spatial hot-spots, giving feedback on temperature readings or increased levels of ethylene. A second application arises in the movement of military troops or police while securing a building. Having information stating that a fellow troop is around the corner could ensure correct action is taken. However, these systems may have drastic consequences if, when relied upon, the WSN security is breached. An officer would be given data that the noise he or she heard was a friendly, but in reality was a foe. In the case of the fire prevention, a node failure could occur preventing the data to be transmitted that a fire was imminent. Clearly, security is a concern. Some applications have higher security requirements than others. However, all WSNs are susceptible to attacks and should be avoided.

Security has long been a trade-off between performance, ease of use, and minimizing risk. Finding the tipping point for a security model and implementation remains a challenge. Compounded with the use of wireless communications and limited energy reserves, a proper design has become even more difficult. A large amount of work has been put into the development of encryption schemes and secure communications, but little has been done analyzing the different data storage techniques [1]. In this paper, the three major data dissemination methods (external, local, and data-centric storage) are analyzed for security implications and issues. For each security concern, the methods are compared to determine which is more susceptible.

In the remainder of this paper, Section 2 gives a deeper background into WSNs and dissemination methods. This provides a better understanding of the issues involved, and why they are issues at all. Related research work in the area of WSN security is given in Section 3. In Section 4, the methods are analyzed against a set of common security issues, deciding on which method

is more adept to overcoming the threat, followed by concluding remarks.

## II. Background

The network is composed of sensor nodes, a gateway, and a host, as shown in Figure 1. The host, often connected via the internet or some other network, issues tasks for the nodes to perform. Connected to the same network is the gateway node, which bridges the two networks. Sample tasks may be to report every occurrence of a radiation reading above a given level, or may be to report the humidity every 10 minutes. Also, multiple tasks may be issued to the same node. After the tasks have been distributed, sensor readings begin. Depending on the dissemination method, the new data may be stored within the network or sent straight out to the host. In the case of network storage, data is kept on the nodes until the host makes a query for the data. Using proper routing algorithms, the data is sent from the nodes to the host where the data is then analyzed.

To better understand the reasons for different dissemination methods, as well as security implications, a brief overview of WSN characteristics is in order. This follows into a discussion of the types of storage methods. Answers to why one method may be better over the others is covered, giving their respective strengths and weaknesses.

### A. WSN Characteristics

Current technology has continually been shrinking the size of sensor nodes, also called motes, from brick-sized devices. A long-term goal for nodes is to be so compact and lightweight that they may be scattered into the atmosphere where they would stay afloat. In this case, smaller is better. With this smaller size, however, comes a cost. There is less space to house processors, memory banks, antennas, and, most importantly, batteries. Without a consistent power source, the network will not function. Having limitations on energy puts limitations on other aspects of the node, such as transmission strength and processing power.

Wireless communications is costly in terms of energy usage due to transmissions. From [6], radio transmissions can be nearly 2.5 times that of the processor. Luckily, the architecture of a WSN is built to service a highly dense network of sensors [1]. This allows the transmission range for any particular node to be lower as neighboring nodes are closer. Each node is then defined a certain transmission strength as to allow communication to a small amount of nodes to conserve energy, giving a small circumference of coverage. In order for a node to relay information to the ultimate sink of the data, communication hops must be made from node to node. This is shown in Figure 1.

Lastly, processing power remains meager, often in the proximity of 5 to 20 MHz [1]. This is a very limiting factor from the viewpoint of security. In mobile computing, on-the-fly encryption for communications is practically a non-issue. The performance drop cannot be realized. Sensor nodes, on the other hand, are less equipped to perform such encryption/decryption techniques. Not only would the process take up more time and possibly cause the node to miss an event or reading, but this causes more strain on the power source. As noted later, many researchers are actively looking to provide a new encryption and decryption method that is both meaningful and cost-effective in terms of power consumption.

### B. Data Dissemination Methods

One of the major aspects of WSNs is how the sensory data is stored or disseminated [2]. These different concepts, explained below, are primarily designed around transmission energy conservation and storage capacity.

*1) External Storage:* Each time a sensor reading is taken, the data is transmitted back to the host, as seen in Figure 1. No storage is required of the nodes which lowers the cost both monetarily and by power consumption for extra components. However, this method has a major drawback: every piece of data requires a transmission and, thus, an amount of energy. Even worse is the trickle effect these transmission have due to the "hopping" nature of communications. In the case of a node with three other nodes in between it and the gateway, three total transmissions are needed for just one piece of data. It is obvious this does not scale well.

External storage is best applied in changing environments where data is not generated at a rapid pace. In a changing, mobile environment, the nodes may not always know where the data-centric storage node (see below) is located, or it may become infeasible to do so. Transmitting directly to the gateway avoids this issue. For a task where readings are taken at large time intervals, the increased transmission count becomes less of an issue.

*2) Local Storage:* One way to combat the transmission energy expenditure of external storage is to keep a small amount of data stored on the node which makes the data. An example network is shown in Figure 2. A node may collect a series of readings which can then be queried for by the host. In the case of a query for the maximum reading, the node can process its data and return a single report instead of the entire series. The cost of this method is two-fold.

Since the host does not know where the data is that is required, the entire network is flooded with queries. This lends local storage methods to periodic measurement sampling. When every node is queried, every node with have a meaningful response. Secondly, every node must implement a data store which will consume more power and increase costs.

*3) Data-centric Storage:* The third major method to data dissemination comes as a marriage of external and

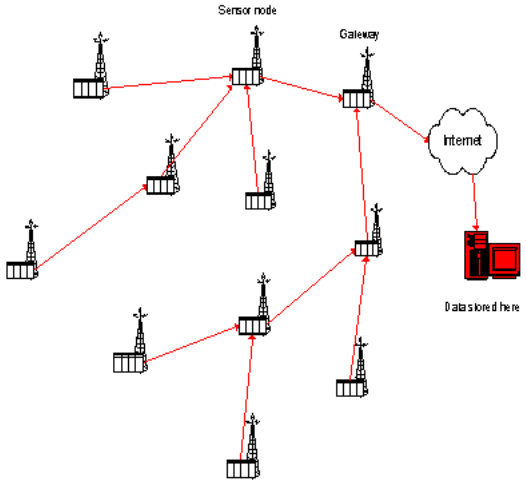


Fig. 1. Network using an external sink for data.

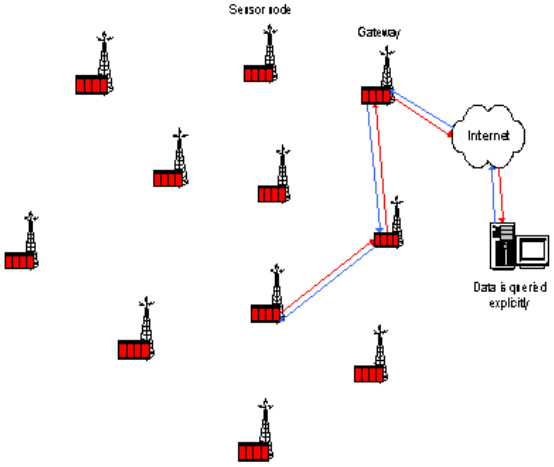


Fig. 2. Network in which all nodes have storage elements. Data is retrieved via queries.

local storage. Instead of all nodes sending data directly out to the host, all readings for a particular task are saved at the same node, seen in Figure 3. Data-centric storage allows the host to easily retrieve the desired data by querying one specific node. To allow for redundancy and resiliency, other nodes can be used to either mirror the data or segment the network.

This method is the most widely adopted. It is not without flaws, however. All data for a particular task are transmitted to the storage node. This node becomes a hot-spot of activity which impacts not only its own energy reserves, but those of surrounding nodes. This primary storage node also induces a single point of failure. Backups and fail-overs are required which add another layer of complexity.

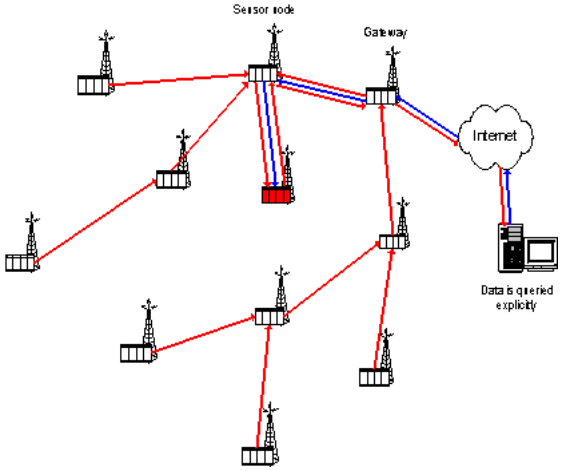


Fig. 3. Network in which one node stores data for a given task. Data is retrieved via queries.

## III. Security Implications

There are many aspects to security at the data dissemination level. In the following subsections, many of the largest security threats are defined and analyzed. Each type of threat falls into one of the three major security views: confidentiality, integrity, and availability [7]. Confidentiality can be defined as the ability of a system to service requests for only allowed nodes. An unauthorized node should not be able to access the data or tasks of another node. Integrity states that an unauthorized node should not be able to change data or tasks of another node. Availability is the ability of the node or network to function, particularly under attack. Each dissemination method is then compared to each other for each particular threat. A simple relative comparative metric is used where a 0 is no impact and a 5 is a high impact. Table 1 gives a summary of the security implications discussed below.

### A. Flooding

The act of flooding in a WSN attacks the network with artificial host queries or tasks or with repetitive data, intent on causing resource exhaustion [1]. These attacks impact the availability of the network. As explained previously, each transmission consumes energy reserves as well as processing time. Query and task floods work at the host level by incorrectly issuing data requests and jobs for the network nodes to perform. Data floods occur when a node sends its sensed data repetitively, beyond its intended operation.

Query flooding only occurs in local and data-centric methods. This is because external storage sends its data immediately, without the host explicitly requesting it. Local storage nodes are impacted the most. When a query is initiated, the host does not know where the requested data resides. All that is known is that each node holds its own data. Thus, queries are sent to the entire network. A

| Characteristic | C,I,A | External Storage | Local Storage | Data-Centric Storage |
|---|---|---|---|---|
| Query Flooding | A | 0 | 5 | 2 |
| Task Flooding | A | 5 | 5 | 5 |
| Data Flooding | A | 5 | 1 | 3 |
| Gateway Compromise | C,I,A | 5 | 4 | 4 |
| Single Node Data Loss | I,A | 1 | 2 | 4 |
| Unauthorized Reads | C | 0 | 4 | 2 |
| Node Movement, Removal, and Replacement | C,I,A | 1 | 1 | 3 |
| Injection of Invalid Data | I | Depends on implementation. | | |

TABLE I
SUMMARY OF SECURITY IMPLICATIONS FOR DATA DISSEMINATION METHODS.

query flood exploits this characteristic by causing large transmission counts to occur. Data-centric networks work differently in response to queries. Here, the host and/or network knows where the requested data is stored - at a node previously specified. Upon the instance of a query, the destination node is determined and the query is sent to that node only. In some cases, such as large sparse networks, multiple nodes are reserved for storing the data for a particular task. Each additional node increases the impact of query flooding.

Task flooding impacts all dissemination methods. In this case, a task is distributed throughout the network, or to a particular subset of nodes, repetitively. All methods use the transmission of tasks in order to collect new types of data. Regardless of the storage method, task flooding affects networks the same.

Networks that sample data frequently or have high occurrences of events are susceptible to energy and storage exhaustion. External storage has the highest impact as each piece of data is sent out immediately to the host, which causes a snowball effect in transmissions. Local storage is the least concerned in this attack as the data resides locally until queried for. Transmission energy is not under question here, but rather storage reserves. Depending on the task, each node may require a new memory location for each piece of data, which is very limited. This problem is worsened by data-centric methods as one node must hold all the data from the network for a given task. Memory can easily be exhausted. However, nodes intended to hold the data are generally more robust with more memory capacity. Transmission counts are increased in data-centric storage because the sensed data must travel along the network to reach the storage node. In most cases, the distance between these nodes are minimal, but not always.

### *B. Compromise at the Gateway*

All communication between the host and the sensor network must pass through the gateway. Loss of this node will break all ties between the host and network. Compromises may affect confidentiality and integrity, but more importantly availability. In confidentiality and integrity compromises, all dissemination methods are affected equally. Availability compromises are slightly different, but in an important way. In local and data-centric storage, data is stored until requested - a link between the sensing nodes and the host does not necessarily incur data loss. Since queries are not initiated, energy from transmissions is not lost, but memory will eventually run out. External storage relies on the host-network link via the gateway. When broken, data will ultimately be dropped as there is no sink for it. Transmissions are affected unless the nodes can "learn" about gateway failures and enter into a dormant phase.

### *C. Data Loss Due to a Single Node Failure*

Node failure is inevitable. The larger the network, the higher the probability a node will fail. Failure here is considered to be when a node ceases to collect and compile data. The impact of node failure on data loss depends highly on the node's job and its previously compiled data. Networks using external storage are least affected. Only after data has been gathered and before the data is sent can data loss occur, and only this newly gathered data would be lost. Local storage is similar to external storage in these regards. The difference is that the data previously gathered at the failed node would be lost. The impact would be closely tied with the frequency of queries. Data-centric storage has the highest potential for large data loss as all data for a task is stored at one node. Other node failures will be similar to external storage nodes. However, a clever attack could monitor the transmission patterns and determine the node with highest source and sink transmission and thus determine the storage node. Resiliency to this attack may come in the form of redundancy - using multiple nodes to store the data. This type of threat falls into the category of data integrity and availability.

### *D. Unauthorized Reads*

Confidentiality comes into play in applications where a leak of information is undesirable. In the case of temperature readings in anticipation of wildfires, unauthorized reading of the data stored in a network may not be a concern. Military and commercial WSNs may collect sensitive data, and thus unauthorized reading is a concern.

A read is performed by manually retrieving information from a node's data store. This is in contrast to reads via transmission, which are out of the scope of this paper. Clearly, external storage is not susceptible at all to this threat as nodes do not storage any data to read. Local storage nodes that are compromised can have their data read, but only their data. Depending on the placement of the node with respect to the frequency of sampling and events, the impact may be small or great. Data-centric storage has the same problem as noted above, all task-related data is stored at a single node. If that node is compromised, all data is susceptible to unauthorized reads, whereas other nodes are no confidentiality risk.

### *E. Node Movement, Removal, and Replacement*

Physically moving, removing, or replacing a node is the most fundamental storage threat in WSNs. Without a node to sense and store data, the network will not be able to operate at full capacity. (For simplicity, it is assumed nodes do not have nearby redundant nodes in which either node can be used.)

Moving a node can have the same impact as removal - the network does not know where the node is. For external storage, this is not a concern. Data need not be stored anywhere except the host and relies on other nodes purely for transmission hops. Local storage operates in the same fashion. Data-centric storage stands to have a higher risk due to its centralized storage. As with single node failures, data may be lost if the storage node is moved or removed, thus causing integrity loss. Replacing a node can have a drastically different impact on data-centric storage. Using a node specifically altered to perform as the attacker wishes (relay information, alter data, etc.), the entire object of the network will be compromised. This may lead to availability and confidentiality breaches.

### *F. Injection of Invalid Data*

Altering the sensing environment for a node or network is a real concern for all storage methods. Consider the temperature sensing application one more time. If an attacker decided to artificially increase the temperature around a single node, monitoring agencies could instantiate an area-wide exodus to prevent loss of life. All storage methods are equally affected, but the extent of the damage is determinant on the application. The application may range from reporting a maximum value, a single event, an average, or so on.

## IV. Future Work

The next stage in this work would be to extend the analysis to use experimental networks to gather security information using a quantitative measurement in lieu of a comparative assessment. This could be performed with usage of the ns2 simulator or by physical nodes. This analysis can be extended further to analyze the security implications of particular implementations, such as the Geographic Hash Table (GHT) method for data-centric dissemination.

Analyze security methods used in the gateway could also be performed as gateway compromization affects all storage methods. The gateway should incur some form of encryption and resiliency measures.

## V. Conclusions

This paper has covered a few of the main security implications that threaten wireless sensor networks and their data dissemination methods. The range of these threats is often determinant of the application of the network. Also, the type of dissemination method varies with application. Viewing the methods without regard to the application shows that external storage is the most secure. However, security is always a trade-off for something else. In this case, external storage requires the most energy expenditure. Local storage ranks second but requires all nodes to have a data storage element, driving up the cost of the unit. Local storage could be improved upon by aggregating data during query responses to further drop the energy impact without compromising security. Data-centric storage is the most widely deployed method as it has the lowest impact on energy with minimal storage, but was found to be the least secure.

## References


[1] Y. Wang, G. Attebury, and B. Ramamurthy, "A Survey of Security Issues in Wireless Sensor Networks," *IEEE Communications Surveys and Tutorials*, Vol. 8, No. 2, 2006.

[2] M. Jeedigunta and X. He, "Evaluation of Data Dissemination Methods in Wireless Sensor Networks," *STAR Lab Internal Report*, 2008.

[3] S. Ravi and A. Raghunathan, "Security in Embedded Systems: Design Challenges," *ACM Transactions on Embedded Computing Systems*, Vol. 3, No. 3, 2004.

[4] S. Ratnasamy, D. Estrin, R. Govindan, B. Karp, S. Shenker, L. Yin, and F. Yu, "Data-centric Storage in Sensornets," *SIGCOM 2002*, 2002.

[5] S. Ganeriwal and M. Srivastava, "Trustworthy Sensor Networks: Issues, Challenges & Solutions," *NESL Technical Report*, UCLA, 2004.

[6] Jason Hill, Robert Szewczyk, Alec Woo, Seth Hollar, David Culler, Kristofer Pister, "System Architecture Directions for Networked Sensors," UC Berkeley, 2000.

[7] Matt Bishop, "Computer Security: Art and Science," Addison-Wesley, 2003.