Integrity Checker for Wireless Sensor Networks in Health Care Applications

Annarita Giani
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA. 94720
Email: agiani@eecs.berkeley.edu

Tanya Roosta
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA. 94720
Email: roosta@eecs.berkeley.edu

Shankar Sastry
Department of Electrical Engineering and Computer Sciences
University of California, Berkeley
Berkeley, CA. 94720
Email: sastry@eecs.berkeley.edu

***Abstract*—Wireless sensor networks (WSN) for health care systems are used to transmit large amount of data collected from several physiological and environmental sensors. Because the information regarding the health of an individual is highly sensitive, it must be kept private and secure. It is of paramount importance to defend the network against any illegal access, as well as malicious insertion of data that would alter the integrity of the entire system. In this paper, we propose solutions to ensure robustness, integrity, and privacy of sensor networks in health care systems. In addition, we define new metrics for determining the integrity of the sensory data. These metrics are defined based on specific characteristics of the health care systems.**

I. Introduction

The new health care solutions utilize ad-hoc networking principles and wireless communication technology of sensor networks to provide continuous and non-intrusive health monitoring regardless of a patient's or a care-giver's location and activity. Examples include, CustoMed [1] and the deployed system proposed in [2]. In addition, sensors are used in the hospital setting to allow doctors, nurses and other care-givers to continuously monitor the vital signs and status of their patients. Given the time critical nature of health care interventions, Wireless Sensor Networks (WSN) in medical applications must have well-defined reliability and trustworthiness metrics. The issues regarding the laws and policies surrounding privacy of health care information has been discussed in detail in [3]. The main focus of the research presented in this paper is to develop tools to defend the health care system against malicious intrusions, such as the insertion of malicious data, which would result in compromising the integrity of the entire system and the privacy of its users. In fact, an important aspect of preserving integrity relates to the system itself rather than only to data items.

The design requirements for a medical sensor network depend greatly on the specific application and deployment environment. In [4] the authors identify several characteristics that are shared by almost all medical sensor networks. These include: wearable sensor platforms, reliable communications, multiple receivers, device mobility, device lifetime and security.

Sensor networks pose unique challenges in terms of designing security mechanisms, specifically because of power, computation and communication constraints of the individual sensors. As a result, security techniques used in traditional networks cannot be applied directly to the sensor networks paradigm. Moreover, given the unattended nature of the sensor network and the broadcast communication medium, the security threat is much higher than in traditional networks. There are many research efforts in securing wireless networks for health care systems. However, many of these efforts do not consider the issue of data integrity and system robustness. For example, [5] examines the utility of physiological parameters for generating cryptographic parameters in order to secure the communication in a wearable body area sensor network. Our approach is fundamentally different from previous research in that it focuses on designing Intrusion Detection System (IDS) and Integrity Monitoring System (IMS) specifically for WSN in health care system.

The main contributions of this paper are: characterizing the types of data alteration, proposing metrics for determining the severity of attacks on the data integrity and reliability, and proposing architectures for an IDS and an IMS. The rest of the paper is organized as follows. Section II describes the design issues of IDS and IMS

PERVASIVEHEALTH 2008, 30 Jan - 1 Feb. Tampere, Finland

DOI 10.4108/ICST.PERVASIVEHEALTH2008.2528

for WSN health care systems. Section III describes the components of our proposed solution, and Section IV explains the details of the integrated IDS and IMS system designed to protect the integrity and privacy of the health care sensor network.

II. Intrusion Detection System for Health Care WSN

Assuring system robustness in the context of this work refers to: assuring that the system is performing well (availability, reliability and accountability), and all of the system components, including any software or hardware, are not faulty.

Traditionally, cryptography has been used as the first line of defense to ensure data integrity. However, cryptographic schemes works under the assumption that an attacker is not able to gain access to hardware. This assumption fails in the sensor network paradigm since these networks are generally deployed and left unattended. If an adversary captures a sensor, he can easily extract the cryptographic primitives and keys as well as exploit the shortcomings of the software implementation [6]. Once the adversary has the cryptographic key, he is able to access the sensory data in order to modify or to exploit it. This motivates the need for an IDS, i.e. an application-level module that is able to detect and to take the right countermeasures against attacks.

Data can be altered or corrupted in two ways: *Maliciously*, when an attacker, who has accessed the cryptographic keys, modifies the information in the system, or *Incidentally*, when a hardware malfunction result in incorrect data. Regardless of the cause of corruption, the IDS must be able to detect abnormal data.

A distinct characteristics of wireless sensor networks in health care applications is that the entire system is composed of completely orthogonal entities; sensor nodes, people and the network used to transmit data. We must consider each of these components, their behavior, and the possible threats posed by each of them in order to best design techniques to ensure integrity and reliability of the entire system.

III. Proposed Components

In this section we discuss the four main components of our system and conclude with ways to integrate them.

A. Data Classification

An important issue in designing an IDS is the ability to characterize "good" versus "bad" network data. Once the models for acceptable and unacceptable traffic have been built, the incoming observations will be compared to these models and classified as acceptable or anomalous. When suspicious observations are detected, an alert is triggered. For example, the detection of anomalous data should prevent drug delivering actuators from engaging and request prompt investigation by a member of the medical staff. Possible models to classify the incoming data are:

Value range: For each sensor output the highest and the lowest possible values are registered and if the incoming data are outside this range, the data is considered malicious. This very simple model protects the system against naíve attackers that do not have any knowledge of the meaning of the data they are corrupting.

Behavior in time: Values of each observation usually follow certain patterns. Therefore, some behaviors in time can not be possible. For example, a person's heart beats 70 times in a minute, and that is the pattern one expects to see when monitoring a patient. Therefore, if we observe a sporadic pattern of heart beat that does not follow the pattern we are used to, this might raise a flag in the system. The model of how each observation value evolves in time can be built by considering the spatial and temporal behavior of the monitored object using the gathered data. It is worth mentioning that depending on the application of the health care system, the model of acceptable behavior in time might change. For example, if the system is monitoring a cardiovascular patient, then a sporadic heart beat might signal a heart attack instead of a malicious attack. Therefore, the model for behavior in time is dependent on its application.

Feature similarity: Observations from different sensors are not completely independent of one another. For example, heart rate and activity level of a person are highly correlated. We need to model this correlation among various observed signals so that the system is capable of determining unacceptable correlations and triggering an alert if an anomaly is detected. These types of models are the most effective to defend the network against the most skilled attacker. He/she can modify the range in a way that models based on behaviors do not trigger any alerts but it less probable that an attacker has the knowledge and the power to modify all the sensor readings to maintain their correlations acceptable. The implementation challenge here is to identify the interdependence among sensor readings. This requires support from medical doctors that know how sensory observations correlate.

B. System Criticality Metric

WSN for health care applications can be categorized according to their criticality. This does not depend on the network itself, but exclusively on how the network is used. For example, the use of sensors to monitor leisure physical activities is less critical than the use of sensor devices for heart rate monitoring of patients in a cardiovascular center. In the latter case, threats to integrity and reliability of the system translate into serious consequences to the patient's health. If sensors are connected to actuators the result can be quite harmful.

Some of the dimensions that must be considered are: threat to the life or health of a person, loss of an individual's reputation, or financial hardship. The complete analysis of all the characteristics of the system implementation will require a long time and a great deal of effort.

C. Relevance of Observations

Depending on the application at hand, the observations have "high", "moderate", or "low" relevance to the goals of the system. This relevance is related to the sensitivity of each observation. We define the sensitivity as the magnitude of the change to the status of the system given the added observation.

While criticality is a characteristic of the entire application, observation relevance is a characteristic of each output given the particular application.

D. Robust Inference

The robust inference engine comprises the final component of an IDS. The task of this engine is to make recommendations to the system administrator and actuators so that these entities are capable of making proper decisions. Once the traffic anomalies are detected and identified as malicious tampering or hardware malfunction, the robust inference engine uses these results to provide suggestions to the system.

IV. The Integrated System

In order to implement the aforementioned systems, we need algorithms that are capable of tracking anomalies and intrusions based on streams of observations. Each possible scenario is a complex processes that can be detected and tracked. The large volume of data collected by WSN is a combination of informative data and background noise from the environment or sensor hardware. Powerful correlation engines must be used to detect activities given noisy observations.

In order to describe the correlation engines, we must first describe activities to be detected through Hidden Discrete Event System Models (HDESM). HDESM are discrete event dynamical system models whose underlying internal state space is not directly observable. The distribution of an observation of a HDESM is typically given by a probability distribution conditioned on the hidden state of the system. Two computational problems arise in this framework. The first problem is determining the most likely process-to-observation association, called the Discrete Source Separation Problem. The second problem is concerned with determining the "best" assignment of events to process instances given a sequence of observations and an HDESM process.

Our challenge is to find solutions to the above mentioned problems. These problems are not solved by traditional information retrieval and database query approaches. In our system design, we plan to use the Markov Chain Monte Carlo Data Association (MCMCDA) [7] method and the Process Query System (PQS) Engine [8]. These powerful tracking techniques compute rated hypothesis of consistent tracks given the observed data. These algorithms have proven to be very successful and robust in many challenging applications [9], [10], [11], [12].

The MCMCDA methodology is an efficient real-time algorithm that solves the data association problem and is capable of initiating and terminating a varying number of tracks. MCMCDA is suitable for sensor networks since it operates with no or incomplete classification information. Process Query System is a powerful software front-end to a database or a real time sensing infrastructure, that allows users to define processes at a high level of abstraction and submit process definition as queries. Missed detection and disambiguation of multiple processes are handled within the PQS kernel. The system uses a library of hidden state sequence estimation algorithms, such as Viterbi-type algorithm for Hidden Markov Models, and Kalman-type algorithms to evaluate state sequences.

The MCMCDA hypothesis generation is based on both current and past observations whereas PQS is a Multiple Hypothesis Testing (MHT) type of tracker that uses recursive filtering techniques to estimate the current state of each track as a function of the current observation being assigned to it, and the previous estimate. As a consequence, MCMCDA does not need to maintain multiple hypotheses, but it needs to reconsider all the observations in order to compute a new hypothesis. PQS, on the other hand, updates the estimate of each track

much more efficiently. However, PQS needs to consider several possible associations of observations to existing tracks which could lead to a potentially geometric growth of the hypotheses set. Given these differences, a future research direction is to compare the two protocols in terms of their applicability to the health care systems.

V. Conclusion and Future Work

WSN for health care systems transmit a large volume of data collected from several bodily sensors. Given that the information regarding the health of an individual is highly sensitive and critical, it must be kept private and secure. Consequently, it is important to equip the system with mechanisms that would prevent unauthorized agents from acquiring or tampering sensitive data that is transmitted over the network and stored in dedicated repositories.

In this paper, we proposed security solutions that included developing Intrusion Detection and Integrity Monitoring Systems by defining the relevant metrics and describing the components of the system. Finally, we discussed MCMCDA and PQS algorithms, which are the two possibilities for developing an Intrusion Detection System to track anomalies and intrusions.

This work is an ongoing research project. As part of the future work, we plan to compare MCMCDA and PQS in terms of their speed of convergence and accuracy of the generated hypotheses. This comparison will aid in determining which algorithm is a better fit for the health care system application. In addition, we plan to define a state space model based on the real health care data, which is available to us from a test-bed, that can be used by MCMCDA and PQS.

References

[1] R. Jafari, "Medical embedded systems," Ph.D. dissertation, University of California, Los Angeles, 2006.

[2] A. Milenkovic, C. Otto, and E. Jovanov, "Wireless sensor networks for personal health monitoring: Issues and an implementation," in *Computer Communications (Special issue: Wireless Sensor Networks: Performance, Reliability, Security, and Beyond)*, 2006.

[3] M. Meingast, T. Roosta, and S. Sastry, "Security and privacy issues with health care information," in *The 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society,*, 2006.

[4] V. Shnayder, B. Chen, K. Lorincz, T. R. F. Fulford Jones, and M. Welsh, "Sensor networks for medical care," in *SenSys '05: Proceedings of the 3rd international conference on Embedded networked sensor systems*, 2005, pp. 314–314.

[5] K.Venkatasubramanian and S.K.S.Gupta, "Security for pervasive health monitoring sensor applications," in *In Proc of 4th International Conference on Intelligent Sensing and Information Processing (ICISIP)*, 2006.

[6] C. Hartung, J. Balasalle, and R. Han, "Node compromise in sensor networks: The need for secure systems," Department of Computer Science University of Colorado at Boulder, Tech. Rep., January 2005.

[7] S. Oh, S. Russell, and S. Sastry, "Markov chain monte carlo data association for general multiple-target tracking problems," in *In Proc. of the 43rd IEEE Conference on Decision and Control, Paradise Island, Bahamas*, 2004.

[8] G. Cybenko, V. H. Berk, V. Crespi, R. S. Gray, and G. Jiang, "An overview of process query systems," in *in Proceedings of the SPIE Vol. 5403*, 2003.

[9] A. Giani, "Detection of attacks on cognitive channels," in *Ph.D Thesis, Thayer School of Engineering, Dartmouth College*, 2006.

[10] I. D. Souza, V. H. Berk, A. Giani, G. Bakos, M. Bates, , and G. V. Cybenko, "Detection of complex cyber attacks," in *in Proceedings of the SPIE, vol 6201*, 2006.

[11] V. Crespi, W. Chung, and A. B. Jordan, "Decentralized sensing and tracking for uav scheduling," in *in Proceedings of the SPIE Vol. 5403*, 2004.

[12] S. Oh and S. Sastry, "An efficient algorithm for tracking multiple maneuvering targets," in *in Proc. of the IEEE International Conference on Decision and Control (CDC), Seville, Spain*, 2005.