Implicit Security Authentication Scheme in Wireless Sensor Networks

Haiguang Chen[1], XinHua Chen[2], Junyu Niu[1]
[1]College of Computer Science
Fudan University
Shanghai, P. R. China, 200433
{hgchen, jyniu}@fudan.edu.cn
2Dept of Electrical Engineering
Hunan Industry Polytechnic
Changsha, P.R. China, 410208
csxinhua@163.com

Abstract

Wireless sensor networks (WSNs) have wide applications due to these sensor nodes ease of deployment.But the security of WSNs is still an important issue. Some existed approaches mainly rely on cryptography to ensure data authentication and integrity. These approaches only address part of the problem of security in WSNs. However, these approaches are not sufficient for the unique characteristics and novel misbehaviors or abnormal behaviors encountered in WSNs. In this paper we introduce an implicit security authentication scheme technique ,which can authenticate the sensor nodes based on actions they would carry out anyway. From our preliminary findings support that this is a meaningful approach, whether used to increase usability or increase security in WSNs.

1. Introduction

Wireless sensor networks (WSNs) have wide range of applications and are used to monitor every type of environment. It provide the system owns with fast and easy access to their data and services anytime and anywhere, especially in remote area such as battlefield, forest ,volcano or otherwise untrusted and even hostile environments. The WSNs work under severe resource constraints, for example, individual sensor node has severely limited computation, storage space, communication and power (battery) resources while operating in settings with great spatial and temporal variability. As in many other network or communications system, the sensor node or the device require protection. However, WSNs are more vulnerable to security attacks than other traditional wired networks and Ad Hoc networks due to their unattended nature. For example, an adversary can physically capture some nodes and use them to inject faulty or false data into the network system disturbing the normal cooperation among these nodes.

The traditional cryptographic and authentication mechanisms used in WSNs, such as TinySec [1], key Session Scheme [2], SPINS [5] , INSENS[6] , SERP[7] and TinyPK [3][10][12], etc., alone cannot be used to solve this problem as internal adversarial nodes will have access to valid cryptographic keys, and on the other hand, these key management schemes only addresses the problem of the node having the valid key to cooperate among the nodes. And furthermore, after a first password entry, it only vouches for the identity of the node or the device, and can't control the behavior because these nodes have the valid keys. Therefore, key scheme doesn't defend against theft and compromise of nodes well, and doesn't address voluntary account sharing at all.

Some contexts use IDSs [4][11][15] to enforce the security of WSNs or using special scheme such as Sybil Attack [8], Wormhole Attack[9] to detect the malicious behavior in system. However, some novel attacks or abnormal behaviors can't be addressed by developing mechanisms that are solely based on cryptography and authentication or IDSs. This is in part because of the uncertainties and lack of control over the physical world and compromised nodes. For example, temporary fault of nodes such as sensor/radio fault or compromised node should be excluded outside the networks. How to detect abnormal behaviors or attacks based on the carry out actions become an important issue.

Yu[13] proposed a machine learning-based intrusion detection system for wireless sensor networks. Every sensor node was equipped an intrusion detection agent (IDA) to detect the intrusions. But the sensor node has limited resource, the IDA can only use part of audit data for LIDC.

Farooqi , A.H. [14]use k-nearest neighbor (KNN) as classifier to detect intrusive attacks in ip multimedia subsystem, the scheme and other scheme[16] need large strage space or heavy compute, these scheme can't used directly for wireless sensor networks.

In this paper, we proposed a scheme which based on machine learning scheme similar as KNN approach

978-0-7695-4258-4/10 $26.00 © 2010 IEEE
DOI 10.1109/MINES.2010.170

to detect the abnormal behaviors of wireless sensor nodes, the approach we refer to it as implicit security authentication. Implicit security authentication can be used to meet the following needs in wireless sensor networks: 1) Used as a secondary factor for security in WSNs, implicit security authentication can augment passwords to achieve higher-assurance security in a cost-effective and user-friendly manner. 2) Used as a primary method of security, implicit security authentication can replace passwords altogether, relieving the nodes from the burden of managing the passwords. 3) A third use of the technology is to provide additional assurance for detect malicious behavior of the nodes. 4) Implicit to detect novel attacks in WSNs.

We based on the behaviors of the nodes to authenticate the node. The main contributions of this paper are listed as follows:

1. Offer a distributed agent-based machine learning model to detect malicious nodes
2. Develop a way to compute the distance away from normal behavior vector in WSNs

The rest of the paper is organized as follows. Section 2 briefly describes the adversarial model in WSNs. Section 3 describes the data and system architecture used in our paper. Section 4 and Section 5 presents authenticate framework and watchdog-based authentication scheme in WSNs, respectively. We give our simulation result and conclusion and point out the further work in section 6 and section 7, respectively.

2. Adversarial Model

In this section, we consider the sensor nodes in the unattendant environment that can be exposed to all kinds of attacks.

One of possible attacks in the WSNs is called node capture. This describes a scenario where an adversary can gain full control over the sensor nodes through direct physical access. The adversary use the validate key to inject false data and disturb the normal cooperate among these nodes. This type of attack is fundamentally different from other kind of attacks.

Except node capture attack, there is some other kinds of attacks in wireless sensor networks, such as radio jamming interferes,exhaustion of battery attacks, selective forwarding, sybil attack, wormholes attack, hello flood, sinkhole,desychronisation attacks, flood attacks, and some unknown attacks or abnormal behaviors. All of these attacks will disturb the normal cooperation among nodes, though the node maybe can do well in one aspect, in our system, the node should exclude from the networks, because the node has done abnormal behaviors in the networks.

3. Data and System Architecture

All the data sources gathered from the behaviors of sensor node used to make implicit security authentication decisions can be grouped into four classes: physical layer data, data link layer data, network layer data and application layer data. Some data may belong to more than one class.

Physical layer data. In this layer, the sensor node provide rich sources of data for implicit authentication. 1)The radio transmission range, it is well known that long distance wireless communication can be expensive,the sensor node usually has the same max transmission range in the system. All data transmit to the neighbor nodes by a node use the radio range always shorter than the max transmission range. 2) Frequency selection is another important data source in physical layer, the entire node use the same radio frequency. 3) Signal detection and propagation.

Data link layer data. The ata link layer is responsible for the multiplexing of data stream, data frame detection, the medium access and error control, the sensor node at this layer, provide important data for implicit security authentication. 1) The error rate, in WSNs, the error rate should below a reasonable value. 2) Collision rate in a node, the max number of packet collisions must be lower than the expected number in the network. 3) Integrity of a packet, the packet in this layer should be integrated used by upper layer. 4) Interval time for packet, the time interval in consecutive packet to the receiver node can not larger or smaller than the allowed limits.5) fairly and efficiently share communication resources.

Network layer data. The network layer provide data for implicit authentication.1) the max number of neighbor nodes, the number neighbors of a node are decrease and can't increase due some of nodes will use out their battery.2) Data delivery, the data delivered by some rules in network. 3) The number of successful delivers data.

Application layer data: This layer include some important protocols to manage the sensor node and the data query.1) the rules related to data aggregation, attribute-based naming, and clustering to the sensor nodes.2)Exchanging data related to the node location.3) Time synchronization of the sensor nodes.4)Querying the sensor network configuration and the status of nodes, and reconfiguring the sensor network.

System architecture.

Watchdog scheme is well-suited to be the trusted third party in charge of making authentication inferences and communicating trust statements, the trusted third party is also possible for nodes simply to

provide data to nodes entrusted with the analysis of data and the making of authentication decisions.

4. Authenticate Framework

We outline the authenticate framework that similar as the machine learning algorithm. We first get data from the node's past behavior which characterizes an individual's behavioral patterns, then we get the normal node behavioral model. This always done before the sensor nodes deploy to the target environment. After getting the node model, the watchdog taking the normal behavioral model and deployed with the sensor node to the target environment. To make a security authentication decision in real-time, a scoring algorithm examines the normal node behavior model and the node's recent behavior, and outputs a score indicating the likelihood that the entity is correct node. The score is used to make an authentication decision. Typically, we can use a threshold to decide whether to accept or reject the node. As described in figure 1.

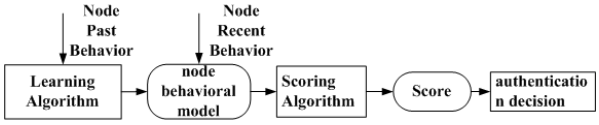

Figure 1: system architecture.

4.1 node normal behavioral model

The node behavioral model can characterize the node's behavioral patterns in normal state. For example, how frequently the node sends data to other nodes, what the radio frequency the node used, the radio transmission range, etc. The node behavioral model may also consider the node's behavioral patterns at different indicators.

We now describe the node behavioral model where we assume independence between different categories of activities. In other words, we assume that the node's one behavior is independent from another behavior. We also assume that some of the node's behaviors depend on the time of staying at networks. For example, one node might have declining neighborhood nodes for some nodes will left network due out of battery.

Let $\xi_1, \xi_2, ...\xi_n$ denote n independent random varibales, we also referred to as n features. For example, ξ_1 =the radio transmission range, ξ_2 =radio frequency, ξ_3 = the error rate, etc.

A node behavioral model (NM) is a product of n probability density functions with the variable t under e ,which t is the length of time the node stayed at the network and e is the event happened in the network.

$$\text{NM: } (p(\xi_1, t \mid e), p(\xi_2, t \mid e), ..., p(\xi_n, t, \mid e))$$

The model is a vector. When we get the node model and the past node behavior which done different event under laboratory, we will use machine learning algorithm described in Figure 1 to establish the node normal behavioral model.

The node behavior can be described by a tuple $(t, e, v_1, v_2, ..., v_n)$, where t denotes the length of time where a node stayed in the network, e is the events and $(v_1, v_2, ..., v_n)$ denote the values of variables $(\xi_1, \xi_2, ..., \xi_n)$. Given the probability density distribution $p(\xi_i, t \mid e)$ and an observed value v_i , the i th scoring function S_i outputs a score s_i for this feature. We can get a series of scores $(t_i, e_j, s_{il1}, s_{il2}, ...s_{iln})$, where $i = 0, 1, 2...$, $j = 1, 2, ..., k$, $l = 1, 2, ..., q$ in different time length t_i and different event e_j with variables $(\xi_1, \xi_2, ..., \xi_n)$. The node behavior vector $(t_i, e_j, s_{il1}, s_{il2}, ...s_{iln})$ might different at event e_j and different time t_i .For example, the average of neighbor nodes to build route path might different at different time. At initial network, the number is bigger than other time due some nodes use up their battery. And at the same time t_i and same event e_j the node behavior vector is different, but the behavior is normal. For example, the error rate might change, but below a certain threshold.

In order to get node normal behavior model(NNM), we use a set of characteristic vectors from these vectors $(t_i, e_j, s_{il1}, s_{il2}, ...s_{iln})$,where i =1,2,.., j =1,2,…k. The characteristic vector is the representative of different event at different time. We use the following approach to get NNM.

For each event j at the time i ,we get three vectors to compose the $NNM(i, j)$:

$$NNM(i,j): \begin{bmatrix} S_{ij}^{Max} = (t_i, e_j, s_{i1}^{Max}, s_{i2}^{Max}, ...s_{in}^{Max}) \\ S_{ij}^{Ave} = (t_i, e_j, s_{i1}^{Ave}, s_{i2}^{Ave}, ...s_{in}^{Ave}) \\ S_{ij}^{Min} = (t_i, e_j, s_{i1}^{Min}, s_{i2}^{Min}, ...s_{in}^{Min}) \end{bmatrix}$$

Where:

$$s_{il}^{Max} = \max\{s_{il} \mid l = 1,2,...q\}$$
$$s_{il}^{Min} = \min\{s_{il} \mid l = 1,2,...q\}$$
$$s_{il}^{Ave} = \frac{\sum_{l=1}^{q} s_{il}}{q}$$

4.2 node behavior authentication approach

Given the node behavior vector $b_{ij} = (t_i, e_j, s_{il1}, s_{il2}, ...s_{iln})$ for n different features at time t_i while doing the event e_j, we call $f_{ij}(b_{ij}, NNM(i,j))$ to compute the score.

$$f_{ij} = \frac{dst(b_{ij}, S_{ij}^{Max}) + dst(b_{ij}, S_{ij}^{Ave}) + dst(b_{ij}, S_{ij}^{Min})}{3}$$

Where $dst(a,b) = \frac{\sum_{i=1}^{n}(a_i \times b_i)}{\sqrt{\sum_{i=1}^{n} a_i^2} \times \sqrt{\sum_{i=1}^{n} b_i^2}}$,

And $a = (a_1, a_2, ..., a_n), b = (b_1, b_2, ..., b_n)$

Every event j at the different time i has different threshold F_{ij}, if $f_{ij} \geq F_{ij}$, then we consider the node behavior is normal, and pass the security authentication. The threshold F_{ij} was defined as follows:

$$F_{ij} = dis(S_{ij}^{Max}, S_{ij}^{Min})$$

5. Watchdog-based Authentication Scheme

We assume that the network consists of a set of scarce resource sensor nodes N. The maximal radio range allowed by power constraints nodes is r. A set of specially equipped nodes M we call watachdogs, which has embedded the NNM before deployed to the target environment. The radio transmission range of watchdog is R ($R > r$). And all the network nodes included watchdogs are deployed randomly in a specific network region of area, A. And a few assumptions listed as follows:

a) Not any new sensor nodes were allowed added into the target WSNs and the sensor nodes are static in the networks.
b) The watchdogs have more power, compute competence, storage space, etc than the normal nodes.
c) The watchdogs work at promiscuous mode listen all the events, packets and messages of nodes around them.
d) There exist malicious sensor nodes, which can do any kind of abnormal behaviors.

Now, we will discuss the watchdog-based implicit security authentication scheme. Our proposed algorithm was divided into the following three phases, a) node behavior vector collection; b)Compute the node behavior score f with the NNM. and c) Broadcast the behavior result.

Data behavior collection:

In this phase, nodes behavior is listened in a promiscuous mode by the watchdog node using a time classify function and the important information is filtered, classified and stored for subsequent computes. The important behavior information tagged with event, time and node ID.

Compute behavior score:

We use Figure 2 to detect the abnormal node behavior.

For each behavior vetctor b_{ij} do
  calculate $f_{ij} = f(b_{ij}, NNM(i,j))$;
  if f_{ij} is greater than threshold F_{ij} then
    b_{ij} is normal;
  else then
    b_{ij} is abnomal;

Figure 2: Pseudo code to evaluate the node behavior

Broadcast behavior result:

After computing the node behavior score, the watchdog use the above rules to determine the node behavior result, and broadcast them to nodes around it. These nodes receive the result and make cooperation decision. How to make cooperation decision is beyond the scope of our paper.

6. Simulation results

We developed our WSNs simulator use Java and Matlab 7, kinds of abnormal behaviors from physical layer to application layer and kinds of error in our simulator.

The watchdog carried the training data set, 4500 normal behavior vectors that build up the NNM, which includes 3 different events at 500 different times.

When the watchdog gets a node behavior vector b_{ij}, then we use the above algorithm which described in Figure 2 to decide the node behavior whether is abnormal.

Figure 3 shows our algorthm.Obviously, our method can detect 30% of the abnormal behaviors with zero false positive rates. And the detection rate reaches 92% rapidly, the false positive rate remains as low as

0.075%. In other words, all the abnormal behavior can be detected at cost of only 3 to 4 false alarms for all the events, cosidering the total number of event at the simulation times.

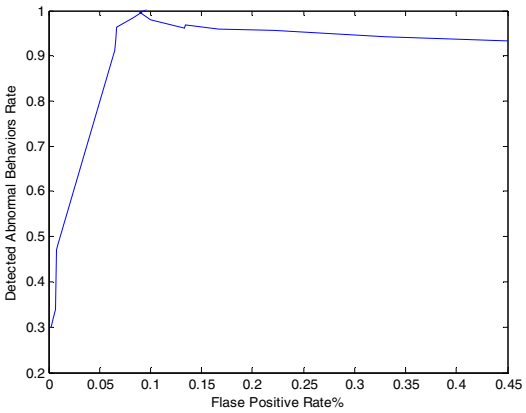

Figure 3: Performance of our method.

7. Conclusion and Further Work

We have proposed a new algorithm to implicit security authentication the node normal behavior and use a scoring scheme to measure the node behavior in the networks. The proposed scheme can increase usability or increase security in WSNs.

As future work, we plan to investigate the following: 1) Research methods to model the relationship between different features (i.e., different activities) in doing a certain event 2) Research methods to model adversarial behavior in wireless sensor networks.

8. Acknowledgment

The research work in this paper was sponsored by the Innovation Programs of Shanghai Municipal Education Commission and the project number is 09YZ154

References

[1]C. Karlof, N. Sastry, and D. Wagner, TinySec: A Link Layer Security Architecture for Wireless Sensor Networks,Proceedings of the Second ACM Conference on Embedded Networked Sensor Systems (SenSys 2004), pages 162-175, November 2004.

[2]Haiguang Chen, Peng Han ,Bo Yu, Chuanshan Gao "A New Kind of Session Keys Based on Message Scheme for Sensor Networks". The Seventeenth Asia Pacific Microwave Conference (APMC 2005) Suzhou , China , Dec. 4-7, 2005

[3]R. Watro, D. Kong, S. F. Cuti, C. Gardiner, C. Lynn, P.Kruus. "TinyPK: Securing Sensor Networks with Public KeyTechnology". In second workshop on Security in Sensor and Ad-hoc Networks, 2004.

[4]W. R. Pires, T. H. P. Figueiredo, H. C. Wong, and A. A. F. Loureiro, Malicious node detection in wireless sensor networks, in 18th Int'l Parallel and Distributed Processing Symp, 2004

[5]A. Perrig, R. Szewczyk, V. Wen, D. Culler, D. Tygar.SPINS: Security Protocols for Sensor Networks. Wireless Networks Journal, September 2002.

[6]J. Deng, R. Han and S. Mishra. The Performance Evaluation of Intrusion-Tolerant Routing in Wireless Sensor Networks. In the Proceedings of IPSN, April, 2003.

[7]S. Ganeriwal, R. Kumar, C. C. Han. S. Lee, M. B. Srivastava. Location & Identity based Secure Event Report Generation for Sensor Networks. NESL Techanical Report, May 2004.

[8]J. Newsome, E. Shi, D. Song and A. Perrig. "The Sybil Attack in Sensor Networks: Analysis and Defenses." In Proceedings of the Third International Symposium on Information Processing in Sensor Networks (IPSN 2004), April 2004.

[9] Y.C. Hu, A. Perrig, and D. B. Johnson, Packet leashes: A defense against wormhole attacks in wireless networks, in Proc of IEEE Infocomm 2003.

[10]An-Ni Shen, Song Guo, and Victor Leung,A Flexible and Efficient Key Distribution Scheme for Renewable Wireless Sensor Networks,EURASIP Journal on Wireless Communications and NetworkingVolume 2009.

[11] Mohi, M. Movaghar, A. Zadeh, P.M.,A Bayesian Game Approach for Preventing DoS Attacks in Wireless Sensor Networks,Communications and Mobile Computing, 2009. CMC '09. 6-8 Jan. 2009,Volume: 3 ,PP507 - 511

[12]Kaiping Xue,Wanxing Xiong,Peilin Hong and Hancheng Lu,NBK: A Novel Neighborhood Based Key Distribution Scheme for Wireless Sensor Networks,pp.175-179, 2009 Fifth International Conference on Networking and Services, 2009

[13]Zhenwei Yu,Jeffrey J. P. Tsai, A Framework of Machine Learning Based Intrusion Detection for Wireless Sensor Networks.Proceedings of the 2008 IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing.PP272-279

[14] Farooqi, A.H. ; Munir, A. ; Intrusion Detection System for IP Multimedia Subsystem using K-Nearest Neighbor classifier,Multitopic Conference, 2008. INMIC 2008. Dec 23-24,2008 ,PP 423 – 428

[15]Di Pietro, Roberto ; Oligeri, Gabriele ; Soriente, Claudio ; Tsudik, Gene ; Intrusion-Resilience in Mobile Unattended WSNs, INFOCOM, 2010, March,14-19 2010,PP1 – 9

[16]Misra,S. Abraham, K.I. Obaidat, M.S. Venkata Krishna, P. Intrusion Detection in Wireless Sensor Networks: The S-Model Learning Automata Approach,1WiMob.2008, Oct.12-14. 2008,PP 603 - 607