國立交通大學

資訊科學與工程研究所

碩 士 論 文

無線感測網路中以統計模型為基礎的異常行為偵測

Detecting Anomalous Behaviors in WSNs with Statistical Learning Model

研 究 生：王朝彥

指導教授：謝續平　教授

中 華 民 國　九 十 六　年 六 月

無線感測網路中以統計學習模型為基礎的異常行為偵測
Detecting Anomalous Behaviors in WSNs with Statistical Learning Model

研 究 生：王朝彥　　Student：Chau-yan Wang

指導教授：謝續平　　Advisor：Shiuhpyng Shieh

國 立 交 通 大 學

資 訊 科 學 與 工 程 研 究 所

碩 士 論 文

A Thesis

Submitted to Institute of Computer Science and Engineering

College of Computer Science

National Chiao Tung University

in partial Fulfillment of the Requirements

for the Degree of

Master

in

Computer Science

June 2007

Hsinchu, Taiwan, Republic of China

中華民國九十六年六月

無線感測網路中以統計學習模型為基礎之異常行為偵測

研究生：王朝彥　　指導教授：謝續平 博士

國立交通大學資訊科學與工程研究所碩士班

摘　　要

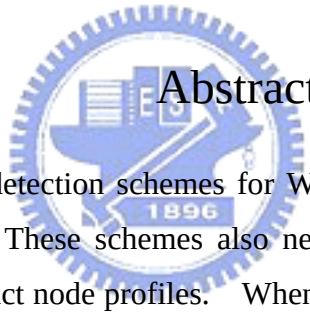
在無線感測網路中，傳統偵測節點異常行為的方法往往需要額外的監控節點來進行監控的動作。這些方法也需要比較冗長的訓練時間來完成對節點建立正常行為模型的動作，根據所建立的行為模型，當有節點行為偏離該行為模型則被認為是異常的行為。這樣的偵測方式通常使用預設的臨界值來辨別異常的活動產生。然而，節點的行為可能會隨著時間的不同而進行改變，所以一個預先設定好的臨界值往往沒辦法精確的分辨網路異常的狀況。在本篇論文中，我們提出一個臨界值設定的方法，該方法藉由結合灰色預測模型及馬可夫模型來建立節點正常行為的模型。除此之外，若節點行為發生改變，該方法可以動態的改變臨界值來適應節點行為的改變。本方法可以容易的使用在無線感測網路中，而不需要額外的監控節點。根據實驗顯示，本方法可以精確而有效的找出無線感測網路中的異常行為。

Detecting Anomalous behaviors in WSNs with Statistical Learning Model

Student: Chauyan Wang Advisor: Shiuhpyng Shieh

Department of Computer Science

National Chiao Tung University



Abstract

Conventional anomaly detection schemes for WSNs require special detection nodes to monitor node behaviors. These schemes also need long training time to model sensor node behaviors and construct node profiles. When a node deviates from its node behavior profile, it is considered as anomaly. In this type of schemes, it is common to use a predetermined threshold to differentiate anomalous activities. However, node behavior may vary over time, and therefore a fixed threshold may not be able to accurately differentiate anomalies. In this paper, we propose a threshold estimation method which combines the Grey Prediction Model and Markov Residual Error Model to model normal node behaviors, and can dynamically adjust the threshold to adapt to the changing behavior of WSNs. Our approach can be easily used in a WSN without the need for special detection nodes. As the experimental results showed, our proposed method can detect anomalous WSN behaviors in a more accurate and effective way than conventional schemes.

Index Terms – sensor networks, detecting anomalous behaviors, dynamic threshold

誌　　謝

這一篇論文可以順利的產生，要感謝的人非常的多。最先也最重要的是我的指導教授 謝續平 教授，因為老師給予的幫助和建議及平時 meeting 不斷的修正我論文的方向，如今才有這篇論文的產生。其次就是 lab 的同學們，仁倩、紀樹、之涯、浩洋、青波，在最後衝刺的階段，大家總是不忘互相勉勵，彼此加油打氣也因此不管再怎麼難熬，還是撐過來了。

另外就是，還要感謝碩一的學弟妹在我們論文繁忙之餘，總是不忘跟我們打打嘴砲緩和一下緊繃壓力，讓我們可以稍稍的喘一口氣。因為有你們，承擔大部份實驗室的計劃，讓我們可以專心的做研究並完成自己的論文。感謝，特別是你們畢業合送的花束以及謝師宴上的小禮物，讓我真的很感動。

除此之外，也要感謝我的家人，因為有你們的支持我才能無後顧之憂的把研究所唸完，謝謝你們。最後我要特別感謝我的女友，淑敏。陪我走過碩班 2 年，這之間因為論文壓力等等常常會感到焦躁不安，而妳的溫柔適時的安撫著我焦慮的心，讓我可以繼續努力下去，謝謝妳。

Table of Contents

List of Figures

1. Introduction

Wireless sensor networks (WSNs) have many applications in many fields [1] [11]. The variety and number of applications are growing on wireless sensor networks, such as military, home safety. WSN requires large number of sensor nodes which collect data from the deployed environment and are for specific tasks. WSN is self-organized with collaboration among the sensor nodes which are tiny devices with limited energy, memory, transmission range, and computation capability. In WSN, base station receives the sensed data from the sensors and does analysis for these sensed data. It is usually a powerful computer with more energy, memory, and is connected to the Internet.

Some applications in wireless sensor network are secure critical [2] [29]. For military applications, WSNs are scattering into an adversary's territory for detecting and tracking the enemy soldiers' and vehicles' directions. Another example is indoor environment surveillance that sensor networks are deployed to detect intruders via a wireless home security system. These applications are more secure critical in WSN and they need some secure mechanisms to make sure security. Sensor nodes in WSN are always unattended and also physically reachable from the outside world, so they are vulnerable in real world for dangerous attacks. Therefore, sensor nodes in WSN must be preventing from these dangerous attacks.

Intrusion prevention measures, such as encryption and authentication methods [9] [10], can be used as first wall in WSN to reduce damage of dangerous attacks, but cannot prevent from inside attackers. For example, encryption and authentication cannot defend against compromised sensor nodes, which carry the private keys and could decrypt the packet by using private keys. No matter how many intrusion prevention mechanisms are used in a network, there still are always some weak aspects that one could exploit to break in. Detecting anomalous behaviors mechanism presents as a second wall of defense and it is a necessity in wireless sensor network.

Wireless sensor network is vulnerable. In order to provide a secure wireless sensor

networks, we need to deploy detection anomalous behaviors techniques for WSN. In our paper, we propose our detection framework for WSNs. We use the Grey Prediction Model [25] [26] [27] [28] and Markov Residual Error Model to model the normal network behaviors in WSN. The method offers upper bound and lower bound thresholds for normal behaviors in WSN and it triggers alarms when anomalies occurred.

The rest of this paper is organized as follows. First, we will provide the background for wireless sensor network, its security issues, and detection techniques in the section 2. We introduce why wireless sensor networks need detecting anomalous behaviors. Second, we give our detection mechanism overview, and then we explain the components detailed design in our intrusion detection system in the section 3. Finally, we will give the analysis and the evaluation of our proposed scheme in section 4, and then we will give the conclusion in section 5.

2. Background

In this section, we first give the security in WSN and its threat model and security requirements. Then, we will introduce the detection mechanism for WSN. WSN could be used in many fields, for example, military, medicine, home safety, environment surveillance, and so on. In WSN, there are hundreds or thousands of small wireless devices to form a wireless network, and these small wireless devices called sensor nodes. These sensor nodes gather the environment's information and then send back to base station. These nodes communicate with each other by using wireless signal and form a wireless network. The opened medium and the environment where sensor deployed make WSN vulnerable.

2.1 Security in WSN

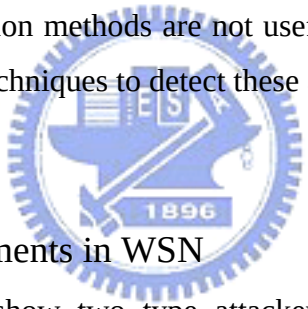
As mentioned before, wireless sensor networks encounter many challenges. Security techniques, for example, authentication and encryption mechanisms used in traditional networks or in MANET cannot be applied directly. Unlike MANET, sensors are often deployed in many inaccessible areas. Besides, sensors interact closely with their physical environments and with people, encountering new security problems from this characteristic. Here we discuss the security included the threat models and security requirements in WSN.

2.1.1 Threat Models

Here we consider two types of attackers: inside attackers and outside attackers. For an outside attacker, the malicious node is not an authorized participant of the sensor network. As the sensor network communicates over a wireless medium, a passive attacker can easily eavesdrop on the network's radio frequency range to steal private or sensitive information from legal node in a WSN. Another type of outside attackers is to disrupt sensor nodes. An attacker can inject useless packets to drain

the receiver's power, or he can capture and physically destroy nodes. A failed node is as same as a disabled node in E. Shi et al. [3]. We could use encryption and authentication methods to prevent from outside attackers.

Different from outside attackers, inside attackers are compromised nodes in the WSN which are the original members in WSN. Compromised nodes seek to disrupt or paralyze the network. A compromised node may be a more powerful device, like laptop or other powerful embedded devices, with more power, large memory size, and powerful computation capability in a WSN. It may run some malicious codes to steal secret information from the sensor network or to disrupt network's normal functions in a WSN. It may have a radio compatible with sensor nodes such that it can communicate with the sensor network. A compromised node can perform arbitrary behavior, which is well known as the Byzantine problem [4]. The encryption and authentication methods are not useful to inside attackers in stead of using the other detection techniques to detect these attacks.



2.1.2 Security Requirements in WSN

In before section, we show two type attackers and then talk about security requirements in WSN. In WSN, authentication is necessary to allow sensor nodes to detect maliciously injected or fake packets. It also allows a node to verify the origin of a packet and ensure *data integrity* in a WSN. Applications in WSN often require data integrity. Although authentication prevents outside attackers from injecting or spoofing packets, it does not solve the problem of inside attackers which are compromised nodes in a WSN. Since a compromised node has the same secret keys of a legitimate node, it can authenticate itself to the network. However, we may need to use some detection techniques to find these compromised nodes and revoke their cryptographic keys in a WSN.

Ensuring the secrecy of sensed data from legal nodes is also important for protecting data from eavesdroppers. We could use encryption mechanisms to

achieve *secrecy requirement*. However, encryption itself is not enough for protecting the privacy of data, as an eavesdropper can do traffic analysis on the overheard cipher-text to get the secrete information, and this can release sensitive information about the data. Another problem is: sensitive data may be released when a compromised node is one endpoint of this communication; or if a globally or group shared key is used. The result is that compromised node can successfully eavesdrop and decrypt the communication between other sensor nodes within its radio range.

Another requirement is providing *availability of network* that the sensor network be functional throughout its lifetime. DoS attacks often result in the loss of availability. In practice, loss of availability may cause serious problems. In an environment monitoring application, loss of availability may cause failure to detect a potential accident and cause some problems. When considering availability in sensor networks, it is important to achieve graceful degradation in the presence of node compromise or benign node failures.

Service integrity is another important requirement. Data aggregation is one of the most important sensor network services because of the data gathering of environment. The goal of secure data aggregation is to obtain a relatively accurate estimate of the real-world quantity being measured, and to be able to detect and reject a reported value that is significantly distorted by corrupted nodes in E. Shi et al. [3].

2.2 Detecting Anomalous Behaviors

Detecting anomalous behaviors is the measure for discovering, analyzing, and reporting unauthorized or damaging network or computer activities to administrator of a network. Detection demands as much information as possible to find anomalies. Intrusion detection techniques can provide digital forensic data to support post-compromise law enforcement actions. It can identify network wrong configurations; improve management and customer understanding of the internet's

inherent hostility.

2.2.1 Architectures of Detection Techniques

Here we do some classifications of the detection techniques for WSNs according to the topology of WSNs. In WSN, the detection techniques are basically belonged to the networks-based detection scheme. According to topology of WSN, two primary intrusion detection models are distributed-based and hierarchical-based detection techniques. In distributed-based detection techniques, every sensor runs its own detection system. In this type, node could cooperate with each other or not. If a sensor cooperates with each other, they create a global detection view and make more secure than non-cooperative nodes in WSNs. But there are some drawbacks in the distributed-based architecture. Nodes may drain their power when always run these detection systems. In hierarchical-based detection techniques, only sink node or cluster-head needs to run the detection system. For example, a cluster-head (CH) could install a detection system and monitor whole cluster. There are some advantages by using hierarchical-based detection scheme: conserving resources prolong the network lifetime, and so on.

No matter in distributed-based or hierarchical-based schemes, observed behavior is characterized in terms of a statistical metric and model. A metric is a random variable x presenting a quantitative measure accumulated over a period. According to D. E. Denning [12], the statistical models may be an operational model, mean and standard deviation model, multivariate model, Markov process model, time series model, etc. The paper [13] analysis the characteristics of the activity graphs, detects and reports violations of the stated policy. It uses a hierarchical reduction scheme for the graph construction, which allows it to scale to large networks. And the photo type of this paper had been successively detected the worm attack.

2.2.3 Detection schemes for WSN

Although anomalous behavior detection techniques are important in WSN, there still few works in this area. There are not many papers working on general detection techniques for wireless sensor networks, exited works are for specific kind of attacks, like [14] [17] [32] [34] [35], and so on; or to particular operations, like routing, localization [5] [15] [16] [23] [30]. Here we simply classify the related works about the detecting anomalous behaviors for WSNs as follow:

- Classify Fault and Attack Scheme
- Identify Legal Neighboring Node Scheme
- Use Pre-Localization Information Scheme
- Reuse Layers' Information Scheme
- Find Most Vulnerable Node Scheme

In the category of classify fault and attack scheme, the paper [18] uses the Hidden Markov Model to model the normal behaviors of sensor network. They need the normal weather information to train the model and find the anomalies when a node reports unexpected information. Another paper [33] also uses Hidden Markov Mode to model the normal behaviors for whole WSN. They think the most important thing is to classify the faults and attacks in WSN. If the detection system judges an error as an attack that makes a false alarm, in order to reduce this situation, authors provide a method by using HMM to classify errors and attacks. After training the model, they could classify what's error and what's attack.

In the category of identify legal neighboring node scheme, these works use some statistical methods or other processes to find out the legal nodes near the detection node. The paper [31] uses some statistical method to find a threshold for normal behaviors in WSN. For example, authors use the packets transition rate to find what normal behavior should be and to classify normal/abnormal behaviors. In order to identify the legal neighboring node in WSN, some works use additional detection node to monitor whole network. The paper [20] uses additional detection nodes as

watch-dog to monitor the traffic near them. They use the predefined rules or some predefined thresholds to distinguish the abnormal/normal nodes. This scheme may install more powerful detection systems to help detecting because of more resources in these detection nodes. In paper [19], authors use the pre-selected rules to find out the anomalies. These rules could be pre-selected according to different network environments. There are also some thresholds in these rules. But these predefined thresholds may not always suit for different network environment.

In the category of use pre-localization information scheme, the paper [24] uses pre-loaded localization information to find some abnormal localization information which returned by sensors. This method uses a probability approach to find out these anomalies. In category of reuse layers' information scheme, the authors think that every layer has some valuable information which could be reuse to find out the anomalies. They reuse layers' information and provide an easy way to identify attacks.

3. Proposed scheme

The previous works of detecting anomalous behaviors for WSN have some problems as follow:

- Machine learning mechanisms need long training time and frequent re-training.
- Some schemes need additional detection nodes to monitor whole network.
- These additional detection nodes have deployments problems.
- Using predefined thresholds may not suit for different network environment.

As mentioned before, some schemes use HMM to identify the errors and attacks. These methods need more training data to model the normal behaviors and long training time in WSN. In the training phase, the detection system may not detect anomalous behaviors that makes whole network insecure. Besides, we need some additional detection nodes to deploy detection system. These additional detection nodes may be another cost for whole application and may not suit for some low-cost applications. These detection nodes also have another problem: the deployment problem. If detection nodes could not cover whole network, it makes the detection effect be low and the network is insecure. Finally, these schemes use predefined thresholds to find out the anomalies. The predefined thresholds could not be use in different network environments.

As mentioned before, different network topologies have different detection system architectures. Our scheme is based on the hierarchical-based topology in WSN and the detection system is installed in cluster-head to monitor whole cluster. In order to save the resources of cluster-head, the detecting scheme must be a light-weight and efficient scheme. Our method is this kind of schemes. Since nodes' behaviors in WSN may change, we need to use a dynamic threshold method and find the upper and lower bound to model node behaviors in WSN. In order to find the upper and lower bound, we must know the

trend of long-term change of node's behaviors and the variety of the every time change. We divide our mechanism into two aspects: the trend of long-term behavior change and the variety of every time change and talk about these two parts.

3.1 Predict Trend of Long-Term Behavior Change

In order to find out the trend of long-term behavior change for nodes in WSN, we use a prediction model to predict the trend of the behavior change. The accuracy and simple computation of the prediction model are the most things because of the constraints of sensors. Here we do some comparisons between prediction models as follow:

Table 1, the comparisons between prediction models

Prediction Model	Training Data	Data type	Training Time
Simple Exponential Method	At least 10 raw data	Equidistant	Short
Holt's Method	At least 10 raw data	Similar trend	Short
Regression Method	At least 15 raw data	Similar trend and regularity	Middle
Causal Regression Method	More than 10 raw data	Mixed type	Long
Box-Jenkins Method	More than 50 raw data	Equidistant	Long
Grey Prediction Method	At Least 4 raw data	Equidistant and Non-equidistant	Short

According to Table 1, we choose the grey prediction model for our detection mechanism because of the characteristics of this model. Grey prediction model only need few input data, short training time, and simple computation which are easily used in WSN. We then express how to model the node's behaviors in WSN.

3.1.1 Estimate Trend of Long-Term Behavior Change

We first show the flow of estimating trend of long-term behavior change in WSN. We model the Feature Set (FS) as follow:

$$X^{(0)}_{featurei} = \left\{x_i^{(0)}(1), x_i^{(0)}(2), x_i^{(0)}(3),..., x_i^{(0)}(t)\right\} \quad where\ x_i^{(0)}(t)\ is\ observation\ of\ \ feature\ i\ at\ time\ t$$

then we build GM(1,1) and predict next ρ steps for feature i as follow:

$$\hat{x}^{(0)}_{featurei}(t+\rho) \quad where\ \rho = 1,2,3,.....$$

The whole process is called trend prediction model of feature i

$$GM_{featurei} : x^{(0)}_{featurei}(t) \rightarrow \hat{x}^{(0)}_{featurei}(t+\rho)$$

After we finished the flow, we then get a long-term behavior change model. We then introduce the steps of building our trend model. The first step is model the observation of feature i:

$$y_i = X^{(0)}_{featurei} = \{x_i^{(0)}(1), x_i^{(0)}(2), x_i^{(0)}(3),....., x_i^{(0)}(t)\} = (x^{(0)}_{featurei}(t)) \qquad (1)$$

The original observations needs to do 1-AGO (the first-order accumulated generation operation) to find out the regularity of the raw data sequence:

$$X^{(1)}_{featurei}(t) = \sum_{q=1}^{t} x_i^{(0)}(q) = (x^{(1)}_{featurei}(t)),\ \text{where}\ i = 1,2,...,h,\ \ t = 1,2,...,N \qquad (2)$$

When we got the 1-AGO sequence, we need to do MEAN operating. This is because we need to generate the parameter matrix to solve the prediction coefficients. We do

MEAN operating as follow:

$$Z^{(1)}_{featurei} = \frac{1}{2} x_i^{(1)}(k) + \frac{1}{2} x_i^{(1)}(k-1) \quad where\ k = 2,3,4,...,t \tag{3}$$

Then, we construct the parameter matrixes for solving prediction coefficients as follow:

$$B = \begin{bmatrix} -Z^{(1)}_{featurei}(2) & 1 \\ -Z^{(1)}_{featurei}(3) & 1 \\ \vdots & \\ -Z^{(1)}_{featurei}(n) & 1 \end{bmatrix}$$

$$y_i = \left[x_i^{(0)}(2), x_i^{(0)}(3), x_i^{(0)}(4), ..., x_i^{(0)}(t)\right]^T \tag{4}$$

By using these two parameter matrixes, we could solve the prediction parameters a, u of prediction model. Here we first show the prediction model as follow:

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u \tag{5}$$

We could see the concept of the grey prediction model is formed as a first order difference equation. We need to find the prediction parameters for every feature i in Feature Set (FS) and then we could use feature i's trend model to find the next change of the feature i. We then solve $\hat{p} = [a,u]^T = (B^T B)^{-1} B^T y_i$ to get time response function,

$$\hat{x}^{(1)}_{featurei}(t+1) = \left(x^{(0)}_{featurei}(1) - \frac{u}{a} \right) e^{-at} + \frac{u}{a} \tag{6}$$

or

$$\hat{x}^{(0)}_{featurei}(t) = -a\left(x^{(0)}_{featurei}(1) - \frac{u}{a} \right) e^{-at} \qquad (7)$$

then we call equation (6) and (7) as a trend of long-term behavior change model for the feature i. Through this model, we could easily find next time step change of the feature i for sensor nodes. We then express how to adjust variety of every time change in next section.

3.2 Variety between Trend and Real Value

After estimated trend model in our environment, we could know the node's behavior how to change in the future and then dynamically change the thresholds. But every time we predict the next change for sensors, there are some variety between the prediction value and real value. We need find a reasonable variety for the next change that reduces the false alarm in our mechanism.

Here we need another prediction model for short-term predicting and use the Markov Model. Markov Model has some characteristics which suit WSN. It is easy to use and has good property for short-term prediction. The simple computation property also makes sensors easy using. We then briefly introduce our variety model and give some definitions. Firstly, we define the residual errors as follow:

$$\text{Re}\,sidual\ Errors = \{e_1, e_2, ..., e_t\} \qquad where\ e_j = x^{(0)}_{featurei}(j) - \hat{x}^{(0)}_{featurei}(j) \qquad (8)$$

The $x^{(0)}_{featurei}(j)$ is real value of feature i at time j, and $\hat{x}^{(0)}_{featurei}(j)$ is the prediction value of feature i at time j. We see the difference between these two values as residual error. When we got the residual errors from past t data, we could model these residual errors as follow:

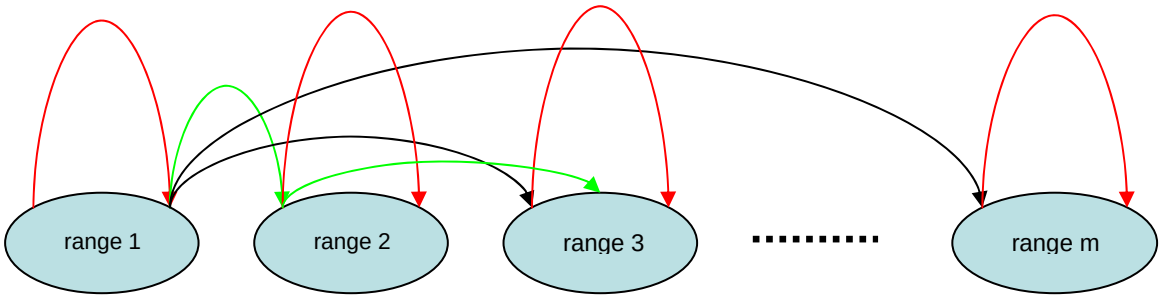


Figure 1, Residual Error Determined Model

Here we first give some definitions of our Residual Error Determined Model as follow:

Initial State (π): the start range of residual errors

State Space (S): the ranges of residual errors

State Transition Matrix (A): $a_{ij}^{(n)} = P(q_{t+n} = s_j \mid q_t = s_i)$

Here we get a feature *i*'s residual error determined model $\lambda = (S, A, \pi)$ to find the variety in next time step. When we got this model, we could easy find how the variety between trend and real value. Then, we add the residual error and trend prediction as the upper bound of normal behavior for sensors. On the other hand, we minus the residual error as the lower bound of normal behavior for sensors. The range is our threshold of feature i for sensors. We will detailed discuss the modeling steps in next section.

3.2.1 Steps of Building Residual Error Determined Model

In this section, we detailed express how to build the model and how to use this model to find the variety. Then, we explain the steps as follow:

- Step 1: Determine the states

  In this step, we need to divide the residual errors which we got

according to past data into m ranges, then we will get

$$State\ i = (e\ , e_y) \quad where\ \ e_x > e_y,\ \ x, y = 1,...,n,\ \ i = 1,...,m$$

- Step 2: Construct the transition matrix A

  According to the states which we determined in previous step, we could assign the errors in residual error set a state. Then, we use this information to build the transition matrix A as follow:

$$A^{(T)} = \begin{bmatrix} p_{11}^{(T)} & p_{12}^{(T)} & p_{13}^{(T)} & \cdots & p_{1m}^{(T)} \\ p_{21}^{(T)} & p_{22}^{(T)} & p_{23}^{(T)} & \cdots & p_{2m}^{(T)} \\ p_{31}^{(T)} & p_{32}^{(T)} & p_{33}^{(T)} & \cdots & p_{3m}^{(T)} \\ \vdots & \vdots & \vdots & & \vdots \\ p_{m1}^{(T)} & p_{m2}^{(T)} & p_{m3}^{(T)} & \cdots & p_{mm}^{(T)} \end{bmatrix} \qquad (9)$$

  where $p_{ij}^{(T)} = \dfrac{q_{ij}^{(T)}}{q_i} \quad i, j = 1,2,3...m$

  the $q_{ij}^{(T)}$ means number of state i transit to state j in time T steps, and q_i means number of the state i in the state table. We could use this transition probability matrix to find the maximum probability and confirm next residual error for next prediction.

- Step 3: Choose the r past varieties to predict r+1 variety

  After finished constructing the transition matrix A, we then do prediction for next variety. Before we do the prediction, we must choose r past varieties and our prediction variety at time r+1 is based on these r past varieties. Here is a trade-off between the computation and precision. More past varieties are and higher precision is, but we will need more computation.

- Step 4: Find the maximum transition probability of state

  We then use the past r varieties to predict r+1 variety. Each variety from past r data far away from the r+1 variety at r step, r-1 step, r-2, step, … , 1 step. According to these distances, we need the r transition matrixes to calculate the probability for each state. We then sum the probabilities of each state to find the maximum one as the prediction variety at time r+1.

- Step 5: Confirm the final variety for next trend prediction

  According to step 4, we decide the state at time r+1. As mentioned before, the state is a error range of the residual errors and we do simple computation to confirm final variety as follow:

$$\hat{e}_{r+1} = \frac{1}{2}(e_x + e_y) \quad x, y = 1...n \tag{10}$$

  When we got the variety for next trend prediction, we add these two values as the upper bound for normal behavior in WSN.

4. Evaluation

We need to build trend model and variety model for our sensors' behaviors. We input the observations of each feature i from Feature Set (FS) to find out the long-term behavior change and variety between prediction value and real value. Both of these models help us dynamically adjust the thresholds to fit the normal behavior changes. In this section we show two phases: one is training phase and the other is testing phase. In training phase, we show flow of constructing the trend model and variety model. In testing phase, we use some testing data to check effect of these two models and differentiate anomalies from testing data.

4.1 Training Phase

We first show the flow of constructing trend and variety models, and then we explain more detail later:

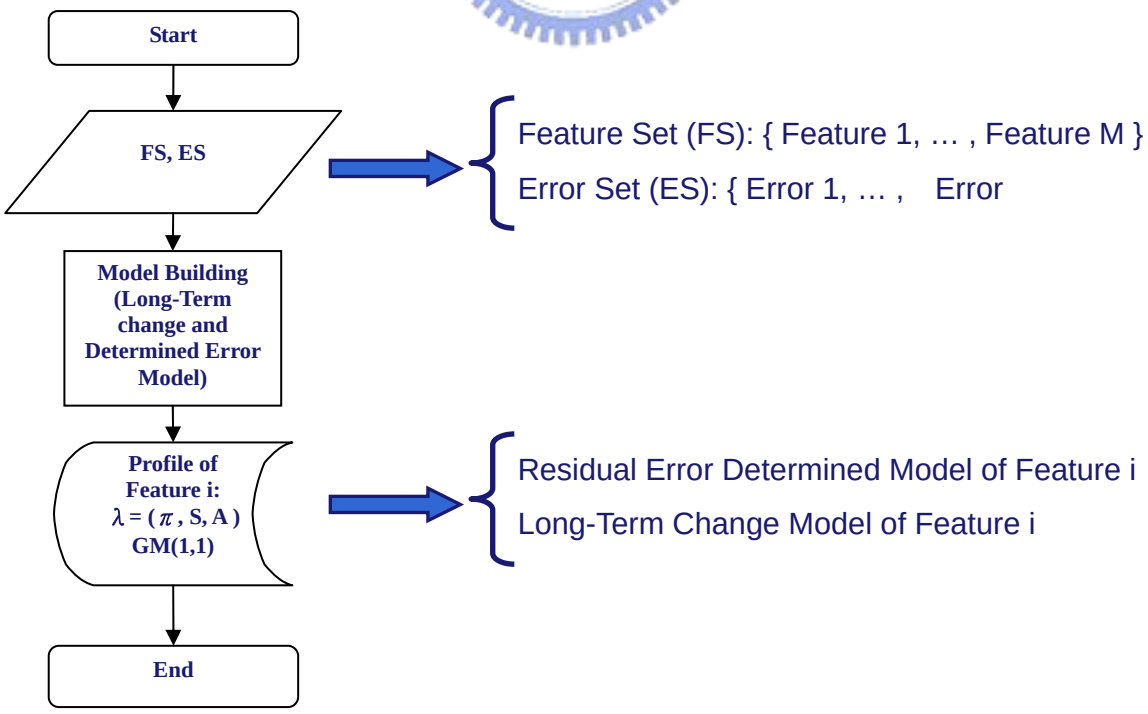


Figure 2, the flow chart of constructing trend and variety models

We input the Feature Set (FS) and Residual Error Set (ES) as raw data to construct trend model and variety model. According to the observations of each feature in feature set, we easy build the trend model. When we built this model, we then predict it to create the residual error set and use the residual error set to build variety model. After finished that, we store these two models in sensors. The sensors could do detecting anomalous behaviors in WSN and then report the anomalies.

4.2 Testing Phase

In previous section, we have already built trend model and variety model. Here we need to test the effect of these two models before we deploy them into sensors. We explain the testing flow and then do evaluation in next section.

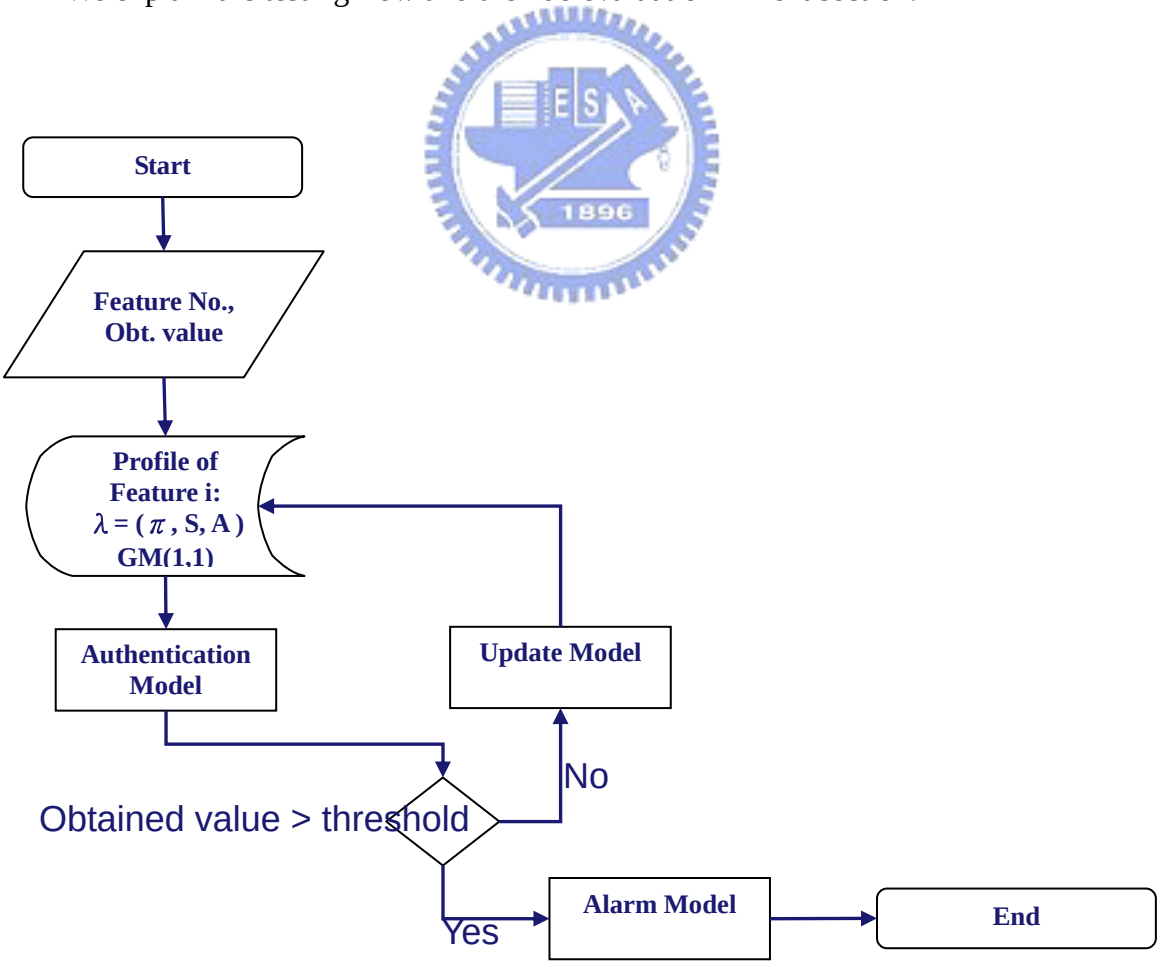


Figure 3, the flow of testing effect of these two models

In this phase, we input the observations of each feature and feature number and then load different feature's trend and variety models. We do two times predicting and then get two prediction values which are trend prediction value and variety prediction value respectively. According to these two prediction values, we could get the upper ad lower bounds of normal behaviors. If the observation doesn't fall into the range, it must be a anomalous behavior. We report a alarm to base station. The detailed steps are as follow:

1. Obtain real value y_i of feature i at time t and input the feature number, observation to do testing.

2. According to the feature number, we load the trend model and variety model of this feature to calculate the trend and variety:

   Trend of long-term behavior change for feature i:

   $$\hat{x}^{(0)}_{featurei}(t) = -a\left(x^{(0)}_{featurei}(1) - \frac{u}{a} \right) e^{-at} \qquad (11)$$

   Variety between trend value and real value:

   $$\lambda = (S, A, \pi) \qquad (12)$$

   These two values decide upper bound and lower bound of normal behavior in WSN.

3. Calculate the threshold for this feature by summing trend value and variety as follow:

   $$threshold\ of\ \ feature\ i\ at\ time\ t = \hat{y}_i = \hat{x}^{(0)}_{featurei}(t) + \hat{e}_t \qquad (13)$$

This threshold of feature i at time t is a temporary value at this time, and next time we will generate a new one for next observation.

4. We compare observation and this threshold. If the observation doesn't fall into this range, we report an alarm to base station. Otherwise, we update the history database of whole network for future constructing the model.

4.3 Result Analysis

In this section, we will discuss our evaluation results of our proposed scheme. We first explain the experiment environment. We applied 21 days measured packet numbers which are accumulated day by day. The sink node accumulates the packets which it got. The first 15 days data are normal network behavior's traffic and are used to train our detection scheme; the last 6 days data are not normal network behavior's traffic and are used to test our scheme. The last 6 days' data contain 2 abnormal data. Here we show the result of the evaluation below:

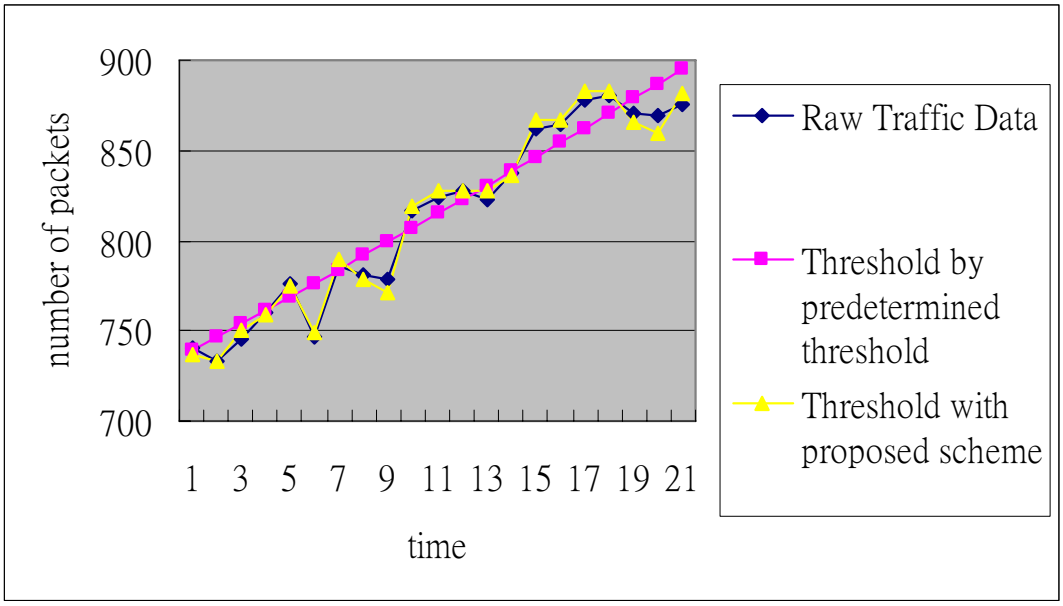


Figure 4, our proposed scheme for detecting anomalies

In this figure, we show the result of predetermined threshold method on measured packet numbers. The blue line is measured packet numbers and the red one is predetermined threshold. The y-axis is number of packets which are bounded between 700 packets and 900 packets and x-axis is measured time of these packet numbers. If we use the red line as a threshold to detect anomalies, we could see there are some normal amount of packets be detected as abnormal activities and some abnormal amount of packets could not be detected. This is why use the predetermined threshold will cause the high false alarm rate because of the predetermined threshold could not dynamically adjust the threshold when the behavior changes in WSNs.

The yellow one in the figure is one which use proposed scheme to model measured packet numbers. We can see that our proposed scheme is more accurate than the predetermined threshold which can dynamically changes the threshold if the behaviors of sensors change. This means we could more easily find out the anomalies in WSN when using the proposed scheme. In our proposed scheme, we use the variety model to promote accuracy of the trend prediction we generated before. The result is good when we test the 6 days' amount of packets.

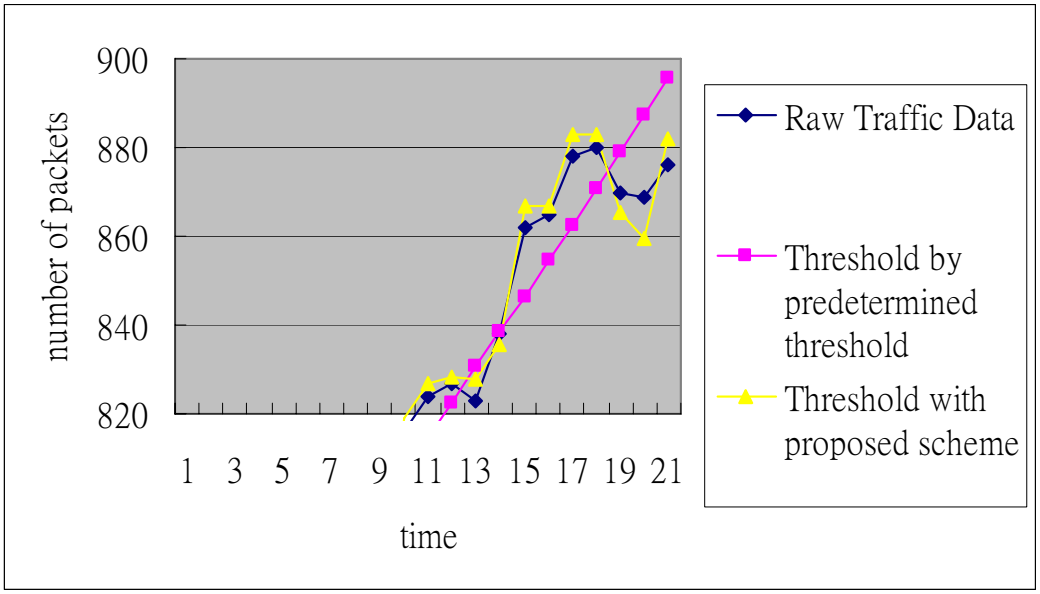
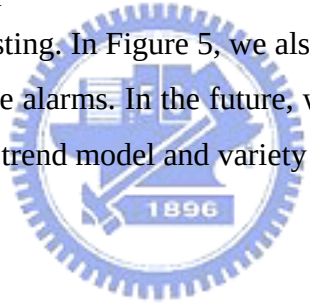


Figure 5, the comparison between the predetermined method and our proposed scheme

As mentioned before, we test 6 days' amount of packets by using our proposed scheme. There are 2 abnormal data in testing data in day 19 and day 20. We can see the predetermined threshold method could not accurately detect the abnormal traffic data from measured traffic data. If we use this method for detecting, there will generate false alarms in day 16, day 17, and day 18 because of the obtained values are great than the thresholds. In day 19 and day 20, the abnormal traffic data will not be detected because of the obtained values less than the thresholds. In our proposed scheme, we can precisely find out the anomalies from mixed normal/abnormal data. We could see the yellow line in Figure 5. In this figure, the yellow line could separate the abnormal network activities and normal network activities precisely. Here we use the yellow line as a threshold to detect anomalies.

5. Conclusion

Here we propose a dynamic adjustment threshold approach included the long-term behavior change model and variety model for wireless sensor networks. When we got an observation value of one feature from the network, we could compare the obtained value with a threshold which is generated by these two models to test whether the obtained value is normal or not in WSN. When the anomalies occurred, the system could report alarms to base station. The advantages of our proposed scheme are that we could dynamically change the thresholds to fit the changes of node's behaviors in the wireless sensor network that improves the precision of detecting anomalous behaviors in WSN environment especially node's behaviors would change over time and also reduce the false alarm generating. In our experimental results, we used measured normal traffic data as an example to train our proposed scheme and then use measured traffic data which included abnormal values for testing. In Figure 5, we also see that the proposed scheme is efficient and can reduce the false alarms. In the future, we will try other statistical model to improve the efficiency of our trend model and variety model.



Reference

1. I.F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A Survey on Sensor Networks,” In IEEE Communication Magazine, August, 2002.
2. A. Perrig, J. Stankovic, and D. Wagner, “Security in Wireless Sensor Networks,” In Communications of the ACM, vol. 47, no. 6, June 2005
3. E. Shi, and A. Perrig, “Designing secure sensor networks,” In IEEE Wireless Communications, Dec 2004
4. L. Lamport, R. Shostak, and M. pease, “The Byzantine Generals Problem,” In ACM Transaction Programming Languages and Systems, vol. 4, no. 3, pages 382-401, July 1982.
5. Chris Karlof, David Wagner, “Secure Routing in Wireless Sensor Networks: Attacks and Countermeasures,” In Sensor Network Protocols and Applications, pages 113-127, May 2003.
6. A. Wood and J. Stankovic, “Denial of service in sensor networks,” In IEEE Computer, 35(10):54-62, October 2002.
7. Tanya Roosta, Shiuhpyng Shieh, Shankar Sastry, “Taxonomy of Security Attacks in Sensor Networks and Countermeasures.”
8. Fei GAO, Jizhou SUN, Zunce WEI, “THE PREDICTION ROLE OF HIDDEN MARKOV MODEL IN INTRUSION DETECTION,” In IEEE CCECE, vol. 2, pages 893-896, May 2003.
9. T. Park and K. G. Shin, “Lisp: A lightweight security protocol for wireless sensor networks,” In Transactions on Embedded Computing Systems, vol. 3, pages 634-660, 2004.
10. W. Du, J. Deng, Y. Han, S. Chen, and P. Varshney, “A Key Management Scheme for Wireless Sensor Networks Using Deployment Knowledge,” In IEEE Infocom, 2004.
11. D. Culler, D. Estrin, M. Srivastava, “Sensor Networks: an Overview,” In IEEE Magazine, Aug 2004.
12. D. E. Denning, “An Intrusion-Detection Model,” In IEEE Transactions on Software Engineering, vol. SE-13, no. 2, pages 222-232, Feb 1987

13. S. Staniford-Chen, S.Cheng, R. Crawford, and M. Dilger, “GRIDS – A Graph Based Intrusion Detection System for Large Networks,” In the 19th National Information Systems Security Conference, 1996.
14. L. Lazos, R. Poovendran, C. Meadows, P. Syverson, and L. W. Chang, “Preventing Wormhole Attacks on Wireless Ad Hoc Networks: A Graph Theoretic Approach,” In IEEE Wireless Communications and Networking Conference (WCNC), 2005.
15. B. Carbunar, L. Loanidis, and C. Nita-Rotaru, “JANUS: Towards Robust and Malicious Resilient Routing in Hybrid Wireless Networks,” In ACM WiSe, Philadelphia, Pennsylvania, USA, October 2004.
16. Y-C. Hu, A. Perrig, and D. B. Johnson, “Rushing Attacks and Defense in Wireless Ad Hoc Network Routing Protocols,” In ACM WiSe, San Diego, Califonia, USA, September 2003.
17. Y-C. Hu, A. Perrig, D.B. Johnson, “Packet Leashes: A Defense against Wormhole Attacks in Wireless Ad Hoc Networks,” In IEEE INFOCOM, 2003.
18. S. S. Doumit, D. P. Agrawal, “Self-organized Critically & Stochastic Learning Based Intrusion Detection System for Wireless Sensor Networks,” In MILCOM, Oct 2003.
19. Ana Paula R. da Silva, Marcelo H. T. Martins, Bruno P. S. Rocha, Antonio A. F. Loureiro, Linnyer B. Ruiz, Hao Chi Wong, “Wireless network security I: Decentralized intrusion detection in wireless sensor networks,” In Proceedings of the 1st ACM international workshop on Quality of service & security in wireless and mobile networks, Oct 2005
20. Rodrigo Roman, Jianying Zhou, Javier Lopez, “Applying intrusion detection systems to wireless sensor networks,” In Consumer Communications and Networking Conference, 2006.
21. Agah, A.; Das, S.K.; Basu, K.; Asadi, M, “Intrusion detection in sensor networks: a non-cooperative game approach,” In Proceedings of Third IEEE International Symposium, pages 343 – 346, 2004
22. P. Techateerawat, A. Jennings, “Energy Efficiency of Intrusion Detection

Systems in Wireless Sensor Networks," In 2006 IEEE/WIC/ACM International Conference, pages 227 – 230, Dec. 2006

23. Q. Fang, J. Gao, L. J Guibas, "Locating and Bypassing. Routing Holes in Sensor Networks," In IEEE INFOCOM'04, March 2004.
24. W. Du, L. Fang, and P. Ning, "LAD: Localization Anomaly Detection for Wireless Sensor Networks," In IPDPS, 2005.
25. Y.-P. Huang and C.-C. Huang, "The integration and application of fuzzy and grey modeling methods," In Fuzzy Sets and Systems 78, pages 107-119, 1996
26. Y.-P. Huang and C.-H. Huang, "Real-valued genetic algorithms for fuzzy grey prediction system," In Fuzzy Sets and Systems 87, pages 265-276,1997
27. Y.-P. Huang and Tai-Min Yu, "The hybrid grey-based models for temperature prediction," In IEEE SMC-B, Vol. 27, No. 2, , pages 284-292, Apr 1997
28. S.-F. Su, C.-B. Lin, Y.-T. Hsu, "A high precision global prediction approach based on local prediction approaches," In IEEE SMC-C, Vol.23, No.4, pages 416-425, Nov. 2002
29. Peyman Kabiri and Ali A. Ghorbani, "Research on Intrusion Detection and Response: A Survey," In International Journal of Network Security, vol. 1, No. 2. pages 84-102, Sep 2005
30. F Anjum, D Subhadrabandhu, S Sarkar, R Shetty, "On optimal placement of intrusion detection modules in sensor networks," In Broadband Networks, 2004
31. I Onat, A Miri, "An intrusion detection system for wireless sensor networks," In Wireless And Mobile Computing, Networking And Communication, 2005
32. ECH Ngai, J Liu, MR Lyu, "On the Intruder Detection for Sinkhole Attack in Wireless Sensor Networks," In IEEE ICC, 2006
33. C Baslie, M Gupta, Z Kalbarczyk, RK Iyer, "An Approach for Detecting and Distinguishing Errors versus Attacks in Sensor Networks," In International Conference on Dependable Systems and Networks, 2006
34. J Newsome, E Shi, D Song, A Perrig, "The Sybil Attack in Sensor Networks: Analysis & Defense," In Information Processing in Sensor Networks, 2004.

35. Waldir Ribeiro Pires, J'unior Thiago, H. de Paula Figueiredo, Hao Chi Wong, Antonio A.F. Loureiro, "Malicious Node Detection inWireless Sensor Networks," In Parallel and Distributed Proceeding Symposium, 2004
36. D. Subhadrabandhu, S. Sarkar, F. Anjum, "Efficacy of Misuse Detection in Adhoc Networks," In Proceeding of the IEEE International Conference on Sensor and Ad hoc Communications and Networks (SECON), Santa Clara, CA, October 4-7, 2004
37. Yi-an Huang, Wenke Lee, "A Cooperative Intrusion Detection System for Ad Hoc Networks," In Proceedings of the 1st ACM workshop on Security of ad hoc and sensor networks, Oct 2003
38. Yongguang Zhang, Wenke Lee, Yi-An Huang, "Intrusion detection techniques for mobile wireless networks," In Wireless Networks, volume 9 Issue 5, Sep 2003