IDMTM: A Novel Intrusion Detection Mechanism Based on Trust Model for Ad Hoc Networks

Furong Wang[1], Chen Huang[1], Jing Zhao[1], Chunming Rong[2]
[1]Dept. of E. I. E, Huazhong University of Science and Technology, Wuhan, Hubei, China
[2] Department of Electrical and Computer Engineering University of Stavanger, 4036 Stavanger, Norway
wangfurong@mail.hust.edu.cn, szo094@hotmail.com, zhaojing_hust@163.com, chunming.rong@uis.no

Abstract

An Ad hoc network is the cooperative engagement of a collection of mobile nodes without the required intervention of any centralized access point or existing infrastructure, so they are vulnerable to many attacks and the security of the network can not be ensued. In this paper, we present a novel Intrusion Detection Mechanism based on the Trust Model (IDMTM) for mobile Ad hoc networks. In IDMTM, we employ two new concepts: "Evidence Chain (EC)" and "Trust Fluctuation (TF)" to accurately evaluate the trust value of a node in the network for judging whether it is malicious or not. Comparing with other Intrusion Detection System, IDMTM can greatly decrease the possibility of false-alarm with by efficiently utilizing the information collected from the local node and the neighboring nodes. Also, IDMTM can efficiently isolate internal malicious nodes from the networks and enhance the security without compromising the performance of the networks.

Keyword: Ad hoc, intrusion, evidence, trust

1. Introduction

In Ad hoc networks, decision-making, key-distribution, routing and forwarding packets are usually decentralized, and many of them depend on the cooperative participation among all nodes. These characteristics of Ad hoc networks allow an adversary to exploit new types of attacks that are designed to destroy the cooperative algorithms such as spoofing, modification of packets and distributed denial of service (DDoS), etc. These attacks can mainly be divided into two parts: external attacks and internal attacks. External attacks can be prevented by many existing methods such as signature, encryption and so on, while there is no efficient technique proposed for internal attacks.

In order to detect the internal malicious nodes, it is necessary to develop an Intrusion Detection System (IDS) which can collect the information about the operation of the network periodically and analyze whether internal malicious nodes exist or not. The conventional IDS mechanisms based on the Neural Network algorithm and the Pattern Matching algorithm can effectively solve the above problems in wired network, but the energy and computing capability of nodes in Ad hoc network can not afford such complicated and computational IDS mechanisms. While in Ad hoc networks, most of the existing IDSs are based on only one aspect of the node's behaviors to judge whether this node is malicious or not. The simplest IDS mechanism is even only depending on the data exchange ratio, which is not accurate enough and always causes false-alarm. False-alarm is considered as one of the main problems that IDS has to face. For example, a conventional IDS might consider the generation of a large number of routing error messages as an form of intrusion, while these messages may be caused by the break of links in fact [1]. Then false-alarm occurs.

In this context, a novel intrusion detection mechanism – IDMTM is proposed. In IDMTM, we employ two new concepts: "Evidence Chain (EC)" and "Trust Fluctuation (TF)" to accurately evaluate the trust value of each node in the Ad hoc network.

For EC, the main malicious behavior forms such as DoS, passive eavesdropping, selfish and so on are regarded as the evidences on judging whether a node is malicious or not. These evidences are composed into an evidence chain. The more a node's behaviors match the components of EC, the more evidences accumulate. Then, we would have sufficient confidence to ensure that this node is a malicious node.

For TF, we give each node in Ad hoc networks a trust value. It is useful to analyze the variety of trust value during a period of time. If the trust value changes sharply, i.e. the sample standard deviation of trust value is great and has the trend to become even larger, we can estimate that the node is being attacked by malicious nodes, or has great probability to change from a normal node to a malicious one.

IDMTM not only effectively utilizes the information collected from the nodes, but also refers to the local neighboring nodes for their recommendations on other nodes to increase the accuracy of the detection of malicious behaviors. So IDMTM can greatly decrease the rate of false-alarm, and pose only a little computational overhead, which would not affect the performance of Ad hoc networks.

The rest of this paper is organized as follows: Related works is given in Section 2. We describe the

1550-445X/08 $25.00 © 2008 IEEE
DOI 10.1109/AINA.2008.124

details of IDMTM in section 3. Section 4 summarizes the results of the simulation on the performance of IDMTM. Finally, section 5 is the conclusion of this paper.

2. Related Work

2.1 IDS Model

Wireless Ad hoc network is subject to different types of security attacks. For example, message bombing, black-hole attack and wormhole attack, rushing attack, and DoS/DDoS attacks. A mechanism such as routing performs best if it can rely on the cooperation of as many hosts as possible. To prevent and detect above attacks and other potential attacks and to secure the communication among wireless Ad hoc network intrusion detection techniques have been studied. An IDS always uses a centralized architecture to collect and analyze audit data to detect unauthorized uses and misuses of computer systems. The IDS can be classified into two categories: signature based intrusion detection, anomaly based intrusion detection. [2,3]

- Anomaly detection: the system knows the user's standard profile and detects deviations from this reference. In [4], Zhang et al. have proposed an anomaly based intrusion detection for Ad hoc networks.
- Misuse detection: relies on the signature of attacks. In signature detection, the intrusion detection decision is formed on the basis of knowledge of a model of the intrusive process and what traces it ought to leave in the observed system. We can define in any and all instances what constitutes legal or illegal behavior, and compare the observed behavior accordingly, e.g. intrusion detection of rushing attacks [5,6] for wormhole detection.

When an intrusion is detected the system may react in different ways, most systems generate an alarm informing the administrator who decided of the reaction to have. A more sophisticated response consists in a corrective behavior (a new rule in a firewall, disconnection of suspicious connections...) to prevent an identical future attack.

2.2 Trust Model

Several works have been pros posed in the literature to deal with the found of trust model.

The distributed trust model, adopted by Abdul-Rahman and Hailes [7], is a decentralized approach for trust management. It uses a recommendation protocol to exchange trust-related information. This model describes how to establish trust relationships. It is applicable to any distributed system and not specifically targeted for Ad hoc networks.

Pretty Good Privacy [8] is an example of system proposed by using a web-of-trust authentication model, and it uses the public key certificate. Another public key approach is similar to PGP in the sense that each user manages its own trust based on direct recommendation.

Hubaux et al., [9] proposed self-organized mobile Ad hoc network, the idea is that each user maintains a local certificate repository. This approach has two drawbacks. First, each user is required to build its local certificate repository before being able to use the system. Second, this approach assumes that trust is transitive.

3. Design of IDMTM

3.1 Overview of the system model

Our model consists of a set of cooperative nodes where each node is running an IDMTM in order to detect an intrusion. The trust value of a node reflects its credit.

When a route request is sent by the source node, nodes receiving this request will check its trust table to find whether the node is trustable. If the sending node can be trusted, the receiving node search its routing table to find a route to the destination, if it can not find the right one, then it will retransmit the route request to its neighbors. So the route can be found at last by retransmitting the route request. The using of IDMTM in Ad hoc network can be illustrated in figure 1.

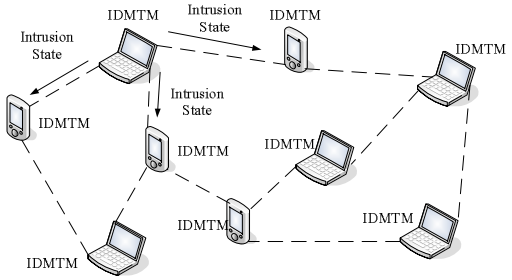

Figure 1. **Cooperative IDMTM in Ad hoc network**

The evaluation of trust value is according to the behaviors of nodes. In IDMTM, the observation on the behaviors of a node is accomplished by all of its neighbors. So when evaluate the trust value of node, we should take reference from the observation results of all the neighboring nodes. According to the trust value, we can categorize each node into different trust levels.

The IDMTM can be divided into two parts: Interior Module and Exterior Module as illustrated in Figure 2.

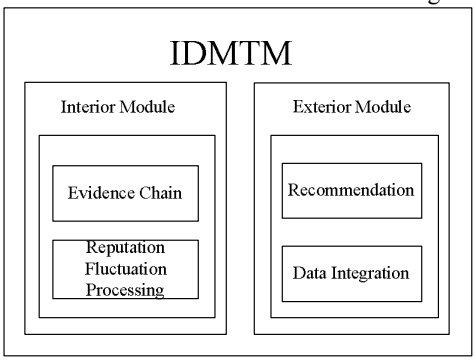

Figure 2. **Architecture of IDMTM**

In the Interior Module, each node monitors the behaviors of a specified node and gives a trust evaluation of it. Interior Module includes "Evidence Chain" and "Trust Fluctuation".

In the Exterior Module, Trust Recommendations (TR) is responsible for negotiating the observation of all the neighboring nodes and Trust Integration (TI) is used to combine all the information together to give an accurate trust evaluation of the node.

3.2 Interior Module

In this module, we mainly focus on the information gathered by the local node; by analyzing them, we can get a local opinion on the investigated node.

3.2.1 Evidence Chain (EC)

In IDMTM, a node captures evidence to quantify the behaviors of other nodes. The evidence is captured by monitoring, acknowledgments and recommendations. The decisions such as whether to send a packet to, or forward a packet on the behalf of another node are dependent on decision policies and evaluation of the captured evidence.

In the law, there is a term called "evidence chain", when all the evidences are collected, they are exit separately and have no law effect. The "evidence chain" is made up of "evidence rings", which describe the relations between the evidence and the truth to be proved. By using these rings, it is enough to prove all the truth.

The "evidence chain" used here is composed of a set of behaviors a malicious node will do to harm the whole network, as in figure 3.

According the source of the attack, we can categorize the attack of Ad hoc networks into external attack and internal attack. The external attack is launched by the nodes without legal identity and certification from CA centre, while internal attacks comes from the compromised nodes which have turned into malicious nodes in the network. Internal attackers may control everything of the compromised nodes, even the private keys or shared secrets with other nodes, so they can pass authentication and generate correct Message Authentication Code for modified or fabricated routing updates. So the internal attacks are more difficult to deal with. By summarizing different types of attack, we can draw out the main malice behaviors of a malicious can be summarized as follows: [10, 2]

- No Forwarding (NF): when receives a packet, the malicious node do not forwarding it to the destination because of selfish or black hole;
- Unusual Traffic Attraction (UTA): generate false routing message that it has the shortest route path to the destination and flood it to the network. Then it can attract all the packets through it and a portion of network resource is wasted by these junk messages and some CPU cycles and memory of nodes are taken up the processing of them.;

Malicious Cheating (MC): inform other nodes that there is something wrong with the present routing which is working well;

- Lack of Error Messages (LEM) : if the route breaks down, the malicious node do not give the error messages to the nodes along the route;
- Frequent Route Updates (FRU) : send the route updates request frequently, it can waste the resources of the network to find a new route;
- Silent Route Change (SRC): modify fields of a routing message, like sequence number and hop counts of AODV, to cause redirection of network traffic;

The behavior of a node changes overtime. A trustworthy node may become malicious later. It is important to monitor the performance of network entities and adjust their trust value timely. When collecting enough monitoring information, we can compare the behaviors with the "evidence chain". If the behaviors of the node match with one or more component of the chain, it means that the node maybe malicious node.

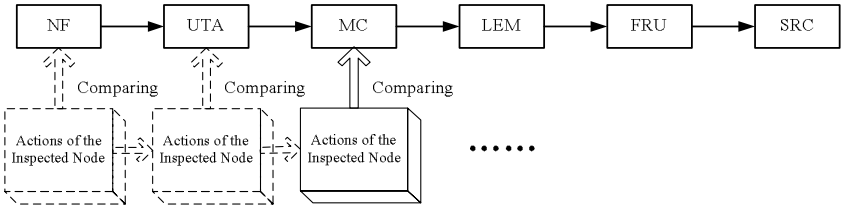

Figure 3. **Illustration of "evidence chain"**

Each intrusion phenomena has been assigned a Weighting Coefficient (WCO) according to its importance in judging the node's character. The trust value of a node is unified to between 0 and 1, so the coefficient must be larger than 1. If the value of WCO is very big, the trust value of a node may decrease sharply when done one more malicious behavior, so the coefficients can not be too large either. The values of the coefficients are suggested as follows:

$$1 < WCO_{SRC} < WCO_{FRU} < WCO_{LEM} < WCO_{LMR} < WCO_{UTA} < WCO_{NF} < 2$$

As the evidence accumulates, we have more confidence to say that it is malicious. To reflect this fact, we choose some Adjusting Coefficients (ACO) to show that the more malicious behaviors a node do, the more trust value will decrease. The values of the adjusting coefficients must stand to the following rule:

$$ACO_1 < ACO_2 < ACO_3 < ACO_4 < ACO_5 << ACO_6$$

Then if the behaviors of the node match n components of the "evidence chain", the Trust Value (TRV) of the node is:

$$TRV = \begin{cases} \dfrac{1}{(WCO_1 + WCO_2 + \cdots + WCO_n) * ACO_n}, n \neq 0 \\ 1, n = 0 \end{cases}$$

If none of the behaviors of the node matches the components of the "evidence chain", the node is considered as the normal node and its trust value is 1. Thus the trust value of a node is between 0 and 1, the larger the value, the high reliability of the node.

For example, the behaviors of the node contain NF, MC and LEM, so $n = 3$. The weighting coefficient for them is defined as:

Table 1. **Weighting Coefficient Table**

Intrusion behavior	NF	UTA	MC	LEM	FRU	SRC
Weighting coefficient	1.8	1.6	1.5	1.4	1.2	1.1

The Adjusting Coefficient Table used here is as follows:

Table 2. **Adjusting Coefficient Table**

Matched number	1	2	3	4	5	6
Adjusting coefficient	1.0	1.5	2.0	3.0	6.0	10.0

When not using the adjusting coefficient, the trust value of the node is:

$$TRV' = \frac{1}{1.8+1.5+1.4} = \frac{1}{4.7} = 0.213$$

By searching the Adjusting Coefficient Table, we can get the adjusting coefficient for the node is $ACO_3 = 2.0$, the trust value of the node is:

$$TRV = \frac{1}{(1.8+1.5+1.4)*2.0} = \frac{1}{4.7*2.0} = \frac{1}{9.4} = 0.106$$

From the above example, we can see clearly that by using the "evidence chain", the trust value can converge to its real value much more quickly and accurately compared to the transitional method.

3.2.2 Trust Fluctuation (TF)

In order to detect whether the node is being attacked by the malicious nodes as early as possible, the sample standard deviation of the trust value is an important quantity to show the tendency. If this value is very large and becomes larger and larger, it means that the node is likely to change into a malicious node.

First we show some definitions of the statistical terms.

The sample measure of spread that is used most often is the sample standard deviation. [11] Let $x_1, x_2 \ldots x_n$ denote sample values. The sample variance, denoted by s^2, is given by

$$s^2 = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n-1}$$

Where $\bar{x}$ is the sample mean which is the numerical average, it is defined by

$$\bar{x} = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}$$

The sample standard deviation, denoted by s, is the positive square root of s^2, i.e.:

$$s = \sqrt{s^2}$$

It is clearly that the sample standard deviation is a measure of variability, large variability in a data set produces relative large sample variance, so it is a measure of the average squared deviation from the mean $\bar{x}$.

In our model, the mean $\bar{x}$ is the average trust value of the node since it joins in the network, and we calculate the sample standard deviation of it every a period of time and compare it with the last item of the sample standard deviation chain. If this value is greater than the last item, it will be store in the chain after the last one; if the value is smaller than the last item, the whole chain will be cleared and the new value will be stored in the first column. If the length of the chain is bigger than a threshold TH, it means the node is on the way to become a malicious node, and some measures must be taken to avoid the bad affects brought by the node.

For example, if the sample standard deviation stored in the chain for the past is 0.1, 0.2, 0.3, 0.4 and the threshold $TH = 5$. If the new sample standard deviation s_5 is 0.5, then the new item will be stored in the chain, we find that the length of the chain is equal to the threshold, so the node is judged as a malicious node and be isolated. If the new sample standard deviation s_5 is 0.3, all the items in the chain are deleted and the new value is store in the first place. The whole process can be illustrated in the following figure:

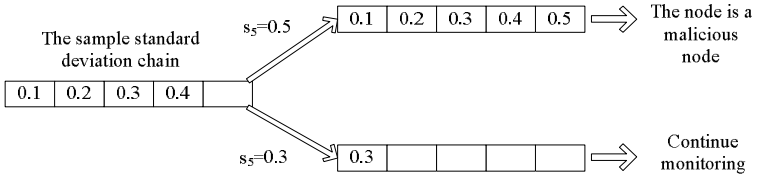

Figure 4. **Trust Fluctuation**

3.3 Exterior Module

When evaluate the trust value of nodes in the network, IDMTM not only effectively utilizes the information collected from the nodes, but also ask its

neighbors for recommendations.

3.3.1 Recommendations

When a node (A) wants to authenticate another node (B), node A first checks its own Trust Table to get a evaluation for node B, but this is not enough. So in our model, node A sends a trust-value-request for node B's trust value to its neighboring nodes. If any of these neighboring nodes does not know node B, this node passes node A's request to its neighboring nodes. So this is a recursive process (measures must be taken to avoid cycle in recommendation) and this trust-value-request may reach a node, node C for example, who knows the trust value for node B. Node C sends a trust-value-reply which carries node B's trust value with the recommender to node A along the reverse path that the trust-value-request travels.

When being asked for the trust value of a specified node, the malicious node may exhibit honest-elicitation by forwarding low recommendations for benign nodes, or high recommendations for colluding malicious nodes. So in order to eliminate the bad effect caused by these untruthful recommendations, we can only accept the recommendations from the nodes whose trust level is equal to or bigger than the querying node (The method to calculate the trust level will be given later). Then the result will be much more consistent with the real situation and less calculation of recommendations will save the power of the node.

The whole process can be illustrated in figure 5, in which different colors means different trust levels.

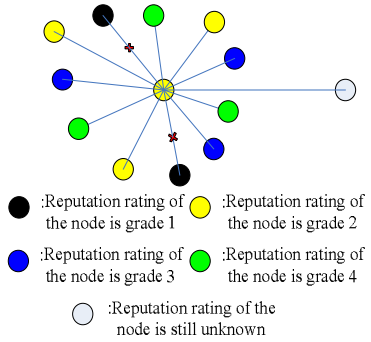

Figure 5. **Negation of Recommendations with Neighbors**

3.3.2 Data Integration

Node A may receive several replies from other nodes then it needs a "Trust Value Evaluation Process" to evaluate node B' trust value. The following function can be used to evaluate the trust value of node B:

$$TRV_B = \frac{\sum_{i=1}^{n} RTRV_i * TRV_i}{\sum_{i=1}^{n} RTRV_i}$$

Where:

TRV_B: The trust value of node B

n: The number of recommendations received

$RTRV_i$: The ith recommend trust value of node B

TRV_i: The trust value of the ith recommend node from the trust-table of node A

Categorizing the trust value of an intrusion to different trust levels can help in decreasing the probability of the false positives and give the response system much more accurate information about the intrusion. Using these trust levels as a guide, the source node can then select a route that meets the security requirements of the message to be transmitted. In many related works, this division is usually linear, but in the real situation, when the node's behaviors match more than one component of the chain, its trust value is usually near to 0. Since the trust value of a node is between 0 and 1, so in our model, the diving intervals near to 0 must be smaller than the intervals for large trust values, just as the shape in the following figure:

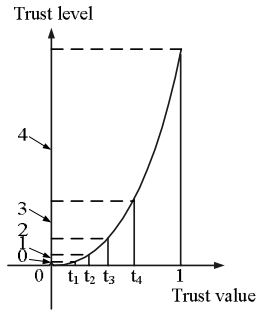

Figure 6. **Mapping Relationship from Trust Value to Trust Level**

Our IDMTM is made up of 5 trust Levels (TLs), each level representing the severity of the attack. We introduce the set of 4 thresholds T to categorize the trust levels, where $T = \{t_1, t_2, t_3, t_4\}$. Then we categorize the trust levels as follows:

$$TL = \begin{cases} 0, 0 < TRV < t_1 \\ r_i, t_{i-1} \le TRV \le t_i; i \in [2,4] \\ 4, t_4 < TRV \le 1 \end{cases}$$

The meaning of each trust level is described in the following table: [12]

Table 3. **The meaning of Each Trust Level**

Trust level	Meaning	Description
0	Compromised	Malicious or known to be compromised
1	Minimum	A low trust level
2	Medium	An average trust level, some what reliable
3	High	A fairly high trust level, considered reliable
4	Highest	An extremely high trust level, considered very reliable

4. Simulation and Discussion

GloMoSim 2.03 simulator is used to simulate our model. We conducted our experiment using Ad hoc On-demand Distance Vector (AODV) protocol as the routing protocol. Our network topology covers the area 1000m by 1000m with 500 mobile nodes. The simulation time is 600 seconds. Our model is compared to a classical IDS mechanism [4]. The classical model consists of set cooperative nodes where each node is running an IDS in order to detect an intrusion. And the coefficients used here are the same as given in the examples.

First, we insert 50 malicious nodes in the network and each time the malicious nodes act different types of malicious behaviors. The simulation result is illustrated in Figure 7. This figure illustrates a comparison between our model and the classical one. It is clear from the figure that as the number of matched malicious behaviors increase, the false-alarm ratio decrease obviously in our model by adopting "evidence chain".

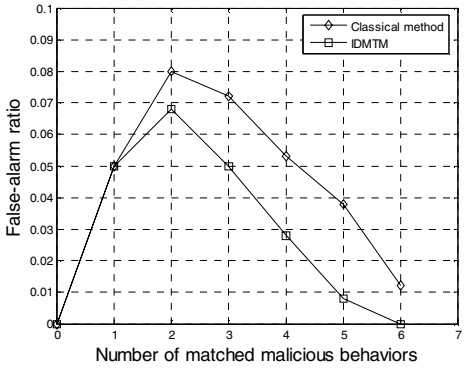

Figure 7. **Simulation Result on EC**

Later we insert 20 malicious nodes in the network, but also with nodes which are attacked by malicious nodes one by one. The following figure shows that our model is much better to detect the nodes which are about to change into the malicious nodes.

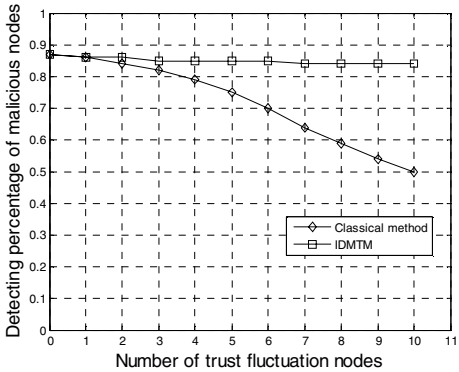

Figure 8. **Simulation Result on TF**

5. Conclusion and Future Work

In this paper, we propose a mechanism to detect malicious node which would cause the decline in the overall performance of the network. By introducing the new concepts "Evidence Chain" and "Trust Fluctuation", the IDMTM system can provide much more accurate judgments on the behaviors of the nodes in the network. The main forms of misbehaviors could be the evidences to accuse a malicious node, and as the evidences accumulate; we have more confidence to ensure a node to be malicious with a low false-alarm. If the trust value of a node is changing sharply, we must consider whether the node is being attacked, or it would turn to be malicious soon. Take some necessary steps in advance as in our model will help in enhancing the security of Ad hoc networks. And the simulation results show that IDMTM is better than some transitional mechanisms when detecting malicious nodes.

However, if a node turns out to be a repenting criminal equivalent that is no longer malicious and has behaved normally for a certain amount of time, some sort of steps to deal with it should be possible. But this is not concluding in this paper and becomes our future work.

Acknowledgements

This work is supported by National Natural Science Foundation of China under Grant No. 60572047 and China Hubei Science & Technology Department through project SBC in 3G CN (2006AA102A04). We thank Furong Wang, Yijun Mo and Chunming Rong for their useful comments on earlier drafts of this paper. The authors would like to thank the anonymous reviewers for their helpful comments to improve the presentation of this paper.

References

[1] Hadi Otrok, Mourad Debbabi, Chadi Assi and Prabir Bhattacharya, "A Cooperative Approach for Analyzing Intrusions in Mobile Ad hoc Networks", *27th International Conference on Distributed Computing Systems Workshops (ICDCSW'07),* 22-29 June 2007 Page(s):86 - 86.

[2] Patrick Albers, Olivier Camp, Jean-Marc Percher, Bernard Jouga, LudovicM´e, and Ricardo Puttini, "Security in Ad Hoc Networks: a General Intrusion Detection Architecture Enhancing Trust Based Approaches".

[3] DaiWei, "Research on Intrusion Detection in Ad Hoc Network", PHD thesis, College of Automation ChongQin University, 2006.5.

[4] Y. Zhang, W. Lee, and Y. Huang. "Intrusion detection technqiues for mobile wireless networks. ACM Wireless Networks", (5):545–556, 2003.

[5] Y.-C. Hu, A. Perrig, and D. B. Johnson. "Rushing attacks and defense in wireless ad hoc network routing protocols", 2002.

[6] Y.-C. Hu, A. Perrig, and D. B. Johnson. "Wormhole attacks in wireless networks". *Selected Areas in*

Communications, IEEE Journal on, 24(2):370–380, Feb. 2006.
[7] Alfarez Abdul-Rahman and Stephen Hailes, "A Distributed Trust Model", *New Paradigms Workshop 1997,* ACM, 1997.
[8] Philip R. Zimmermann, "The official PGP user's guide", MIT Press Cambridge, MA, USA, 1995.
[9] S. Capkun and L. Buttyan and J. Hubaux, "Self-Organized Public-Key Management for Mobile Ad Hoc Networks", ACM *International Workshop on Wireless Security,* WiSe, 2002.
[10] Huaizhi Li, Mukesh Singhal, "A Secure Routing Protocol for Wireless Ad Hoc Networks", *Proceedings of the 39th Hawaii International Conference on System Sciences - 2006*.
[11] Ronald E.Walpole, Raymond H.Myers, Sharon L.Myers, Keying Ye, "Probability & Statistics for Engineeers & Scientists", Prentice Hall, 2002.
[12] Zhaoyu Liu, AnthonyW. Joy, Robert A. Thompson, "A Dynamic Trust Model for Mobile Ad Hoc Networks", Proceedings of the 10th IEEE International Workshop on Future Trends of Distributed Computing Systems (FTDCS'04)