Dynamic Training Intrusion Detection Scheme for Blackhole Attack in MANETs

[1]A.Naveena, K.Rama Linga Reddy, [2] K.Rama Linga Reddy

[1, 2]Electronics and Telematics department, GNITS, Hyderabad, India.

Abstract
Mobile ad hoc network (MANET) is a self-configuring network which is composed of several movable mobile nodes. These mobile nodes communicate with each other without any infrastructure. As wireless ad hoc networks lack an infrastructure, they are exposed to a lot of attacks. This paper analyzes the blackhole attack which is one of the possible attacks in ad hoc networks. In a blackhole attack, a malicious node impersonates a destination node by sending a spoofed route reply packet to a source node that initiates a route discovery. By doing this, the malicious node can deprive the traffic from the source node. In order to prevent this kind of attack, it is crucial to detect the abnormality that occurs during the attack. In conventional schemes, anomaly detection is achieved by defining the normal state from static training data. However, in mobile ad hoc networks where the network topology dynamically changes, such static training method could not be used efficiently. In this paper, we propose an anomaly detection scheme using dynamic training method in which the training data is updated at regular time intervals. The simulation results show the effectiveness of our scheme compared with conventional scheme.

Keywords:
AODV, anomaly detection, blackhole attack, MANET

1. INTRODUCTION

Mobile Ad hoc NETwork (MANET) is a collection of mobile hosts without the required intervention of any existing infrastructure or centralized access point such as a base station. Due to their inherent characteristics of dynamic topology and lack of centralized management security, MANET is vulnerable to various kinds of attacks. Blackhole attack is one of many possible attacks in MANET. In this attack, a malicious node sends a forged Route REPly (RREP) packet to a source node that initiates the route discovery in order to pretend to be a destination node. By comparing the destination sequence number contained in RREP packets when a source node received multiple RREPs, it judges the greatest one as the most recent routing information and selects the route contained in that RREP packet. In case the sequence numbers are equal it selects the route with the smallest hop count. If the attacker spoofed the identity to be the destination node and sends RREP with destination sequence number higher than the real destination node to the source node, the data traffic will flow towards the attacker. Therefore, source and destination nodes are
unable to communicate with each other. In [1], the authors investigated the effect of blackhole attack when movement velocity and a number connection toward the victim node are changed, and proposed the detection technique at the destination node. However, we can effectively avoid the attack for example by selecting the detour route during route reconstruction which achieved by detecting the attack at the source node rather than at the destination node. Thus, taking into account the detection at the source node is indispensable.
Regarding the detection of blackhole attack at the source node, [2, 3] have proposed methods in which still they are using the same training data to define the normal state. However, in MANET where the network state changes frequently, the pre-defined normal state may not accurately reflect the present network state.
In this paper, we use a reactive routing protocol known as Ad hoc On-demand Distance Vector (AODV) routing [4] for analysis of the effect of the blackhole attack when the destination sequence numbers are changed via simulation. Then, we select features in order to define the normal state from the characteristic of blackhole attack [5]. Finally, we present a new training method for high accuracy detection by updating the training data in every given time intervals and adaptively defining the normal state according to the changing network environment.
The rest of this paper is organized as follows. Section II provides the background on the AODV protocol and describes the characteristic of the blackhole attack. Section III analyzes the blackhole attack through simulations. In Section IV, we propose the detection scheme of the attack, and evaluate its effectiveness. Section V concludes the paper.

2. OVERVIEW ON AODV

AODV is a reactive routing protocol in which the network generates routes at the start of communication. Each node has its own sequence number and this number increases when links change. Each node judges whether the channel information is new according to sequence numbers [6]. Fig.1 illustrates the route discovery process in AODV. In

Manuscript received October 5, 2013
Manuscript revised October 20, 2013

this figure, node S is trying to establish a connection to destination D. First, the source node S refers to the route map at the start of communication. In case where there is no route to destination node D, it sends a Route REQuest (RREQ) message using broadcasting. RREQ ID increases one every time node S sends a RREQ. Node A and B which have received RREQ generate and renew the route to its previous hop. They also judge if this is a repeated RREQ. If such RREQ is received, it will be discarded. If A and B has a valid route to the destination D, they send a RREP message to node S. By contrast, in case where the node has no valid route, they send a RREQ using broadcasting. The exchange of route information will be repeated until a RREQ reaches at node D. When node D receives the RREQ, it sends a RREP to node S. When node S receives the RREP, then a route is established. In case a node receives multiple RREPs, it will select a RREP who's the destination sequence number (Dst Seq) is the largest amongst all previously received RREPs. But if Dst Seq were same, it will select the RREP whose hop count is the smallest.

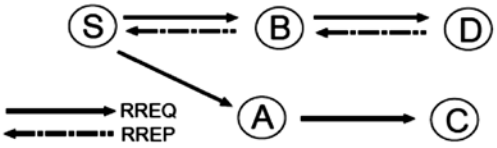


Figure. 1 Route discovery process

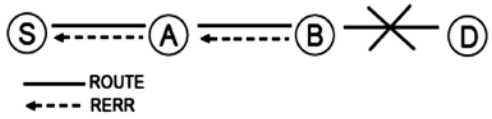


Figure. 2 Transferring route error messages

In Fig. 2, when node B detects disconnection of route, it generates Route ERRor (RERR) messages and puts the invalidated address of node D into list, then sends it to the node A. When node A receives the RERR, it refers to its route map and the current list of RERR messages. If there was a route to destination for node D included in its map, and the next hop in the routing table is a neighboring node B, it invalidates the route and sends a RERR message to node S. In this way, the RERR message can be finally sent to the source node S [7].

A. Description of Blackhole Attack

In AODV, Dst Seq is used to determine the freshness of routing information contained in the message from originating node. When generating a RREP message, a destination node compares its current sequence number and Dst Seq in the RREQ packet plus one, and then selects the larger one as RREP's Dst Seq. Upon receiving a number of RREP, a source node selects the one with greatest Dst Seq in order to construct a route. To succeed in the blackhole attack the attacker must generate its RREP with Dst Seq greater than the Dst Seq of the destination node. It is possible for the attacker to find out Dst Seq of the destination node from the RREQ packet. In general, the attacker can set the value of its RREP's Dst Seq base on the received RREQ's Dst Seq. However, this RREQ's Dst Seq may not present the current Dst Seq of the destination node. Fig. 3 shows an example of the blackhole attack. The value of RREQ and RREP using in the attack are shown in Table 1.

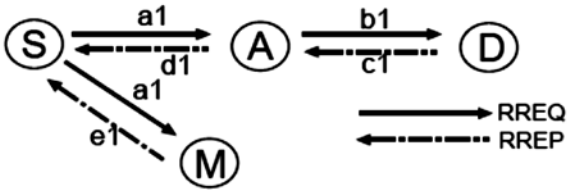


Figure. 3 Blackhole attack

Table 1: Values of RREQ and RREP

	RREQ		RREP		
	a1	b1	c1	d1	e1
IP.Src	S	A	D	A	D (MD)
AODV.Dst	D		D		D (MD)
Dst_Seq	60		61		65
AODV.Src	S		-		-

In Table 1, IP.Src indicates the node which generates or forwards a RREQ or RREP, AODV.Dst indicates the destination node and AODV.Src indicates the source node. Here, we assume that the destination node D has no connections with other nodes. The source node S constructs a route in order to communicate with destination node D. Let the destination node D's Dst Seq that the source node S has is 60. Hence, source node S sets its RREQ (a1) and broadcasts as shown in Table 1. Upon receiving RREQ (a1), node A forwards RREQ (b1) since it is not the destination node. To impersonate the destination node, the attacker M sends spoofed RREP(e1) shown in Table 1 with IP.Src, AODV.Dst the same with D and increased Dst Seq (in this case 65 as) to source node S. At the same time, the destination node D which received RREQ (b1) sends RREP (c1) with Dst Seq incremented by one to node S. Although, the source node S receive two RREP, base on Dst Seq the RREP(e1) from the attacker M is judged to be the most recent routing information and the route to node M is established. As a result, the traffic from the source node to the destination node is deprived by node M.

Next, we consider the case shown in Fig. 4. The value of RREQ and RREP using in Fig. 4 are shown in Table 2. Similar to Fig. 3, source node S attempts to construct a

route to destination node D. However, unlike the environment in Fig. 3, in this case node B, C and E are also constructing a route to D. Therefore, the destination node D's Dst Seq that the source node has is significantly different from the current Dst Seq of node D. Since the most recent Dst Seq from D that node S has is 60, it set RREQ (a2) as shown in Table 2 and broadcasts. Upon receiving RREQ (a2), based on information contained in RREQ (a2) node M sends a spoofed RREP (e2) with Dst Seq 65 the same with previous situation to the source node. Upon receiving RREQ (b2) node D sends RREP (c2) to the source node. However, this time, since node D constructed route with other nodes, we assume that the Dst Seq is increased to 70. Then, this RREP (d2) is forwarded by node A. Upon receiving two RREPs, node S selects the route to destination node D since the Dst Seq of node D is the larger one. As a result, the attack is not succeeded [8].

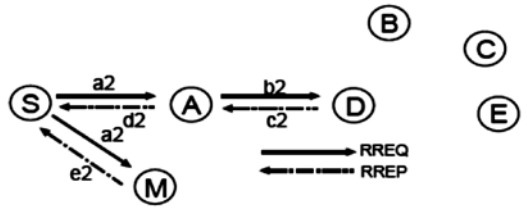


Figure. 4 Blackhole attack in some connections to nodeD

Table 2: Values of RREQ and RREP

	RREQ		RREP		
	a2	b2	c2	d2	e2
IP.Src	S	A	D	A	D (MD)
AODV.Dst	D		D		D (MD)
Dst_Seq	60		70		65
AODV.Src	S		-		-

3. INVESTIGATION OF BLACKHOLE ATTACK

In this section, we investigate the effects of the blackhole attack in MANET using NS2 in our simulation [9]. Depending on the traffic involving in a destination node, its Dst Seq may change. As the recent, in the blackhole attack, the effect of the attack may also change depending on the increased amount of Dst Seq. Here, we specifically investigate the effects of the attack when the number of connections to the destination and the number of connection from the destination are changed.

A. Simulation Environment

For simulation, we set the parameter as shown in Table 3.Random Waypoint Model (RWP) [10] is used as the mobility model of each node. In this model, each node chooses a random destination within the simulation area and a node moves to this destination with a random velocity.

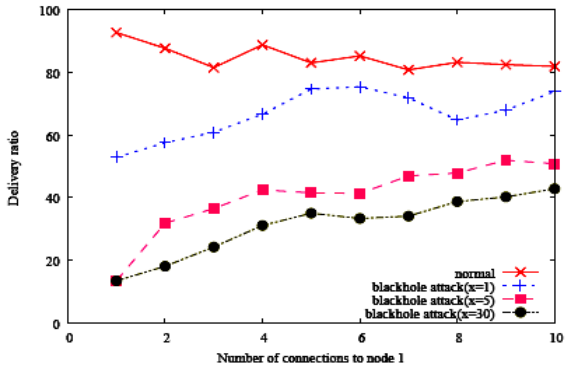
Table 3: Simulation parameters

Simulator	ns-2(ver.2.27)
Simulation time	600(s)
Number of mobile nodes	30
Topology	1000m × 1000m
Transmission Range	250m
Routing Protocol	AODV
Maximum Bandwidth	2Mbps
Traffic	Constant bit rate
Maximum Speed	5(m/s)
pause time	10(s)

Here, we assume that the blackhole attack take place after the attacking node received RREQ for the destination node that it is going to impersonate. Upon receiving RREQ, the attacker set the Dst Seq of RREP to RREQ's Dst Seq + x. Here, x is an integer range form 1 to 30.The node number of each node among 30 nodes in the simulation is given from 0 to 29.

B. Simulation Result of Blackhole attack

First, we investigate the delivery ratio of packet from source node 0 to destination node 1 in case there are connections from other nodes to the destination node. For the
experiment, nodes which are selected randomly from 2 to 28 (except the source node, destination node, and attacking node) generate traffic towards the destination node. Here, we perform experiment by changing the number of nodes generating the traffic from one to nine. This experiment is performed repeatedly five times. Fig. 5 shows the packet delivery ratio from node 0 to node 1. From Fig. 5, we can see that when the number of connection is 1, the more Dst Seq is increased in blackhole attack the more packet delivery ratio drops.
However, when the number of connections increases, the packet ratio increases even when blackhole attack took place.



This is because the destination node's Dst Seq tends to be higher than the attacker's Dst Seq, since attacker set the Dst Seq based on the Dst Seq contained in RREQ coming from the source node. We can see that the more the attacker increases the Dst Seq, the lower the packet delivery rate is.

Next, we investigate the packet delivery ratio from node 0 to node 1 when destination node 1 generates traffic to other nodes. We assume that destination node 1 generates traffic toward other nodes in which their node numbers are randomly selected from 2 to 28 as. The experiment is performed by changing the number of selected nodes from one to ten and this experiment is repeated five times. Fig. 6 shows the packet delivery ratio from node 0 to node 1.

When the number of connections from node 1 increases, in other words, when node 1 initiates more route discoveries to other nodes, Dst Seq tends to be increased. For this reason, the packet delivery ratio increases along with the rising of the number of connections. From these results, we can judge that the Dst Seq of each node change depending on the condition of its traffic.

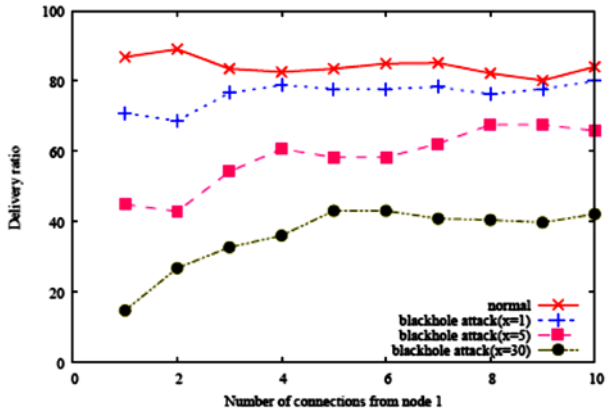


Figure.6 : The delivery ratio versus the number of connec-tions from node 1

4. DETECTING BLACKHOLE ATTACK

A. Feature Selection

To express state of the network at each node, multidimensional feature vector is defined. Each dimension is counted up on every time slot. In order to detect this attack, the destination sequence number is taken into account. In normal state, each node's sequence number changes depending on its traffic conditions. When the number of connections increases the destination sequence number tends to rise, when there are few connections it tends to be increased monotonically. However, when the attack took place, regardless of the environment the sequence number is increased largely. Also, usually the number of sent out RREQ and the number of received RREP is almost the same. From these reasons we use the following features to express the state of the network.

– Number of sent out RREQ messages
– Number of received RREP messages
– The average of difference of Dst Seq in each time slot between the sequence number of RREP message and the one held in the list.

Here, the average of the difference between the Dst Seq in RREQ message and the one held in the list are calculated as follows. When sending or forwarding a RREQ message, each node records the destination IP address and the Dst Seq in its list. When a RREP message is received, the node looks over the list to see if there is a same destination IP address. If it does exist, the difference of Dst Seq is calculated, and this operation is executed for every received RREP message. The average of this difference is finally calculated for each time slot as the feature.

B. Discrimination Module of Anomaly Detection

For the traffic that flow across each node, the network state in time slot i is expressed by three-dimension vector xi = (xi1, xi2, xi3). Here, the groups of normal states are considered to be gathered close in feature space. In contrast, the abnormal state is considered to be the scattering data that deviates from the cluster of normal state. According to this, the distribution of network state is shown. From now, we calculate the Mean vector $\bar{x}^D$ from Equation (1) using training data set D of N time slots.

$$\bar{x}^D = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (1)$$

Next, we calculate the distance from input data sample x to the mean vector $\bar{x}^D$ from Equation (2).

$$d(x) = \|x - \bar{x}^D\|^2 \quad (2)$$

When the distance is larger than the threshold T_h (which means it is out of range as normal traffic), it will be judged as an attack (Equation (3))

$$\begin{cases} d(x) > T_h & : attack \\ d(x) \leq T_h & : normal \end{cases} \quad (3)$$

Here, the projection distance with maximum value is extracted as Th from the learning data set (Equation (4)):

$$T_h = d(x_I), where\ I = \underset{i \quad x_i \in D}{arg\ max}\ d(x_i) \quad (4)$$

Let ΔT0 be the first time interval for a node participating in MANET. By using data collected in this time interval, the initial mean vector is calculated, then the calculated mean vector will be used to detect the attack in the next period time interval ΔT .If the state in ΔT is judged as normal, then the corresponding data set will be used as learning data set. Otherwise, it will be treated as data

including attack and it will be consequently discarded. This way, we keep on learning the normal state of network. The procedure is shown in Fig.7.

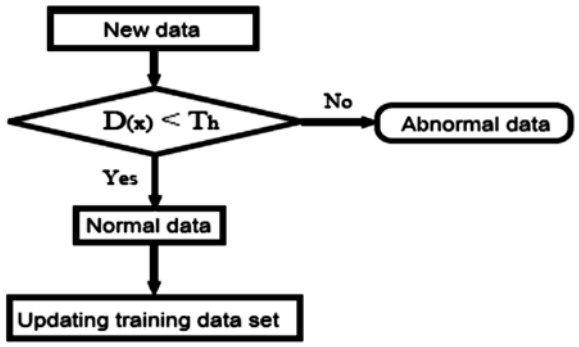


Figure.7 Learning flow chart of proposed method

By doing this, we update the training data set to be used for the next detection. Then, the mean vector which is calculated from this training data set is used for detection of the next data. By repeating this for every time interval ΔT, we can perform anomaly detection which can adapt to MANET environments.

C. Simulation Result

We assume that initial training data set in time interval ΔT0 does not contain attack data, this interval is set to 300(s). Refer to [2, 3], we set the time slot i to be 5 (s).Here, the attacker starts attacking after receiving a RREQ. The Dst Seq of RREP that the attacker sends is equal to the received RREP's Dst Seq increased by x, where x is selected randomly from 5 to 30.From the experiment, the detection rate is shown in Fig. 8, and the false positive rate is shown in Fig. 9. The horizontal axis shows the mobility rate. Here, using initial training data only means that only initial data is used as the training data as in [2, 3].

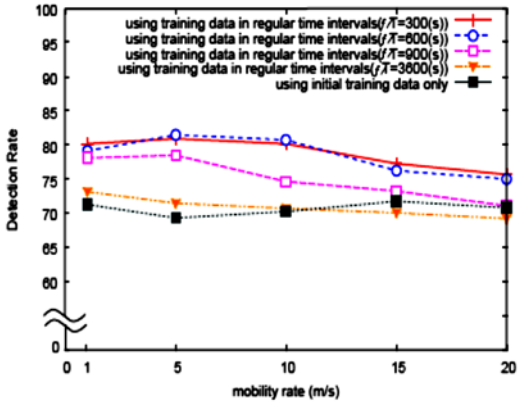


Figure. 8: Detection rate versus mobility rate

From these results, we can see that the detection ac-curacy drops as updating time interval increases. We can also see that it is necessary to shorten the updating interval as the mobility rate become faster. However, the shorter the updating interval is the more processing overhead is needed. Therefore more battery power will be consumed. From these facts, it is necessary to take into account the MANET environment and battery power issue to determine the updating interval. In simulation, even if mobility rate become faster, detection accuracy of the proposed method (ΔT = 300(s)) and (ΔT = 600(s)) are better than the using initial training data only

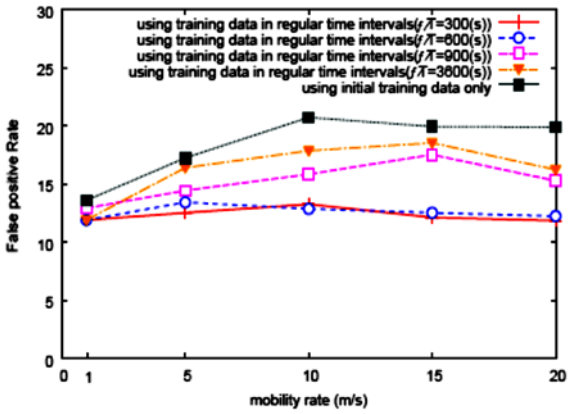


Figure. 9: False positive rate versus mobility rate

However, the detection accuracy of the proposed method degrades when the updating time interval become longer. Comparing the proposed method (ΔT = 600(s)) with using initial training data only, we found that the average detection rate is increased by more than 8% and the average false positive rate is decreased by more than 6%.From this result, we can see that the detection rate and false positive rate has been improved. In the proposed method, by updating the training data it can adapt to the changing environment in MANET, while using initial training data only using only the initial training data can not adapt to the dynamically changing environment. Therefore, we can see that the proposed scheme is effective in anomaly detection [11].

5. CONCLUSION

Blackhole attack is one of the most important security problems in MANET. It is an attack that a malicious node impersonates a destination node by sending forged RREP to a source node that initiates route discovery, and consequently deprives data traffic from the source node. In this paper, we have analyzed the blackhole attack and introduced the feature selection method in order to define

the normal state of the network [12]. We have presented a new detection method based on dynamic learning and updating training data. Through the simulation, our method shows significant effectiveness in detecting the blackhole attack.

REFERENCES

[1] H. Deng, W. Li, and D. P. Agrawal, "Routing security in ad hoc networks," IEEE Communications Magazine, vol. 40, no. 10, pp. 70-75, Oct. 2002.

[2] Y. A. Huang, W. Fan, W. Lee, and P. S. Yu, "Crossfeature analysis for detecting ad-hoc routing anomalies," in The 23rd International Conference on Distributed Computing Systems (ICDCS'03), pp. 478-487, May 2003.

[3] Y. A. Huang and W. Lee, "Attack analysis and detection for ad hoc routing protocols," in The 7th International Symposium on Recent Advances in Intrusion Detection (RAID'04), pp. 125-145, French Riviera, Sept. 2004.

[4] C. E. Perkins, E. M. B. Royer, and S. R. Das, Ad hoc On-Demand Distance Vector (AODV) routing, RFC 3561, July 2003.

[5] M. A. Shurman, S. M. Yoo, and S. Park, "Black hole attack in wireless ad hoc networks," in ACM 42nd Southeast Conference (ACMSE'04), pp. 96-97, Apr.2004.

[6] M.S.Alkatheiri, Jianwei Liu, A.R.Sangi, "AODV routing protocol under several routing attacks in MANETs," 13th International Conference on Communication Technology (ICCT),pp 614-618, Aug.2011.

[7] W. Wang, Y. Lu, and B. K. Bhargava, "On vulnerability and protection of ad hoc on-demand distance vector protocol," in The 10th International Conference on Telecommunications (ICT'03), vol. 1, pp.375-382, French Polynesia, Feb. 2003.

[8] K.A. Jalil,Z.Ahmad,J-L.A.Manan, "Securing Routing Table update in AODV routing protocol," in International Conference on Open Systems (ICOS), pp 116 -121, Sep. 2011.

[9] P.K. Singh, G. Sharma, "An Efficient Prevention of Black Hole Problem in AODV Routing Protocol in MANET," in 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), pp 902 - 906 ,Apr. 2012.

[10] C. Bettstetter, G. Resta, and P. Santi, "The node distribution of the random waypoint mobility model for wireless ad hoc networks," IEEE Transactions on Mobile Computing, vol. 2, no. 3, pp. 257-269, Jul./Sep. 2003.

[11] U. Venkanna, R. Leela Velusamy, "Black hole attack and their counter measure based on trust management in manet:A survey," in 3rd International Conference on Advances in Recent Technologies in Communication and Computing (ARTCom 2011), pp 232 -236, Sep. 2011.

[12] Jiwen Cai, Ping Yi, Jialin Chen, Zhiyang Wang, Ning Liu, "An Adaptive Approach to Detectng Black and Gray Hole attacks in Ad hoc Network," in 24th IEEE International Conference on Advanced Information Networking and Applications (AINA), pp 775 – 780, Apr. 2010.