VANET Routing Replay Attack Detection Research Based on SVM

Qing Gang FAN[1,a], Li WANG[2], Yan Ning CAI[1], Yong Qiang LI[1] and Jing CHEN[1]

[1] *Computer Staff Room, Xi'an Research Institute of High Technology, 710025 Xi'an, China*
[2] *Library, Xi'an Research Institute of High Technology, 710025 Xi'an, China*

Abstract. In the process of establishing routing in VANET, because of the interference of routing request (RREQ) and routing response (RREP), the nodes in the network are busy with looking for routing or establishing routing. The impact on the network performance is extremely serious, such as increasing a lot of network overhead, consuming valuable bandwidth resources. In this paper, the influence of routing replay attack in VANET is studied. Four typical characteristic are extracted by "cross layer" selecting feature vector. The feasibility of VANET routing replay attack IDS based on SVM is verified through simulation experiment.
Keywords. SVM, VANET, Routing Replay Attack

1 ROUTING REPLAY ATTACK

In the communication process based on AODV routing protocol of VANET, establishing routing between nodes is on-demand type. If the source node needs to send a message to the destination node, firstly, the source node will query their own routing table, if the path to the destination node is not exist in the table, source node will be broadcast a route request packet RREQ in the network. Adjacent nodes will query their own routing information table after receiving the RREQ, if the path to the destination node is exist in the table, the RREP message will be sent back directly, passing by relay node to the source node. If not, continued to send RREQ to adjacent nodes, until find the path. The forward routing is established after source node received the RREP. This is the normal communication process based on AODV routing protocol.

In general, a relay node will forward RREP in order to set up the forward routing, when there are malicious nodes in network, malicious nodes would deliberately to interference of communication process, such as not forwarding RREP, makes the forward routing cannot be established. Or not forward RREQ, making the source node can't find the destination node. As shown in figure 1, B is the source node, P is the destination node, B will be broadcast a RREQ message to find the path of the P, in the process of normal communication, relay node, such as C, D, will forward RREQ, and continue to backward forward RREP after find the destination node P. In exceptional circumstances, such as D is a malicious node, after receiving RREP, D does not forward RREP to C, but drops the RREP and give an error, result in C and B can,t receive RREP of P, unable to form traffic route to the destination node P, forming a routing replay attack.

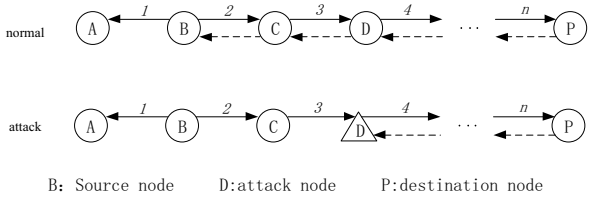

Figure 1. routing replay attack

2 ROUTING REPLAY ATTACK DETECTION METHOD

2.1 Routing replay attack Scenario simulation

This linear model is used to set up routing response team attack scenario in this paper, the specific scene parameter settings are as follows:

Table 1. Routing replay attack Scenario parameter settings

Parameter settings	value
Channel type	wireless channel
Routing protocol	AODV
Interface queue type	Queue/DropTail/PriQueue
Maximum length interface queues	50
Distance between the nodes	70m
Effective communication distance	75m

[a] Qing Gang FAN: fangang3232@126.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

Network interface type	Phy/WirelessPhy
Type of antenna	Antenna/OmniAntenna
Number of vehicles	30
vehicle speed	18m/s
Communications link	3→12 15→20
Frequency of packets	4packets/s
Malicious nodes	Node 10
the size of scene	4000m*500m
simulation time	1000s
Flow type	TCP
CBR packet size	512 bytes
MAC layer protocol type	IEEE 802.11
Wireless propagation mode	Two Ray Ground

By software NS-2, the corresponding simulations with tcl script is run in the normal and routing reply attack scenario, and the corresponding trace files are got. The attacked trace file as shown in figure 2.

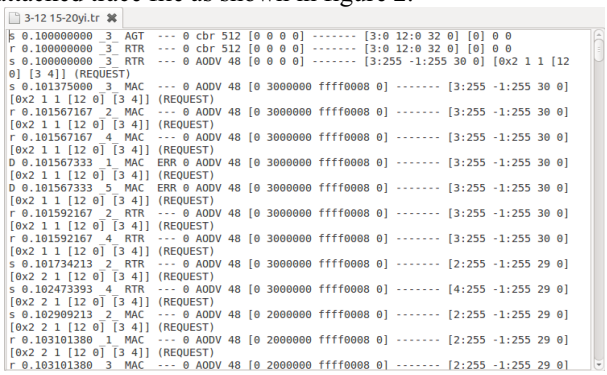

Figure 2. the attacked trace file

2.2 Routing replay attack extracting feature vector

Because the routing request packet and data packets are affected in routing replay attack, the method of cross-layer statistical feature vectors are using in this paper. The send and receive packet statistics and the packets send and receive situation of related to node communication in VANET MAC layer and routing layer are selected as routing replay attack feature vectors. The characteristics of the specific record as shown in table 2.

Table 2. routing replay attack feature

Feature name	Feature description
RREQ_s	Generate RREQ and sends it out
RREQ_r	Receive RREQ
RREQ_d	Drop RREQ
RREQ_f	Forward RREQ
RREP_s	Generate RREP and sends it out
RREP_r	Receive RREP
RREP_d	Drop RREP
RREP_f	Forward RREP
CBR_s	Generate CBR and sends it out
CBR_r	Receive CBR
CBR_d	Drop CBR
CBR_f	Forward CBR

These feature attributes record represents in a transceiver packet types, VANET communication packets record. Such as:

s 2.000000000 _0_ RTR --- 0 AODV 48 [0 0 0 0] ------- [0:255 -1:255 30 0] [0x2 1 1 [19 0] [0 4]] (REQUEST)
r 2.151811992 _7_ RTR --- 0 AODV 44 [0 6000000 8 0] ------- [19:255 0:255 17 5] [0x4 14 [19 4] 10.000000] (REPLY)
f 2.151811992 _5_ RTR --- 0 AODV 44 [0 6000000 8 0] ------- [19:255 0:255 16 4] [0x4 15 [19 4] 10.000000] (REPLY)

The first record as RREQ_s type, show node 0 at the time of 2.0s, broadcast a RREQ packet of node 0 as source node, node 1 as the destination node.

The second record as RREP_r type, show node 7 at the time of 2.151811992s, receive a RREP packet of node 0 as source node, node 19 as the destination node.

The third record as RREP_f type, show node 5 at the time of 2.151811992s, forword a RREP packet of node 0 as source node, node 19 as the destination node.

Reading data from the detailed records, each node in any period of packet types and transceiver can be counted, then formed the available SVM feature vector. For example, by NS - 2 a piece of data simulation to generate the following records:

r 2.016590249 _3_ RTR --- 0 AODV 48 [0 4000000 ffff0008 0] ---- --- [4:255 -1:255 26 0] [0x2 5 1 [19 0] [0 4]] (REQUEST)
r 2.016590249 _5_ RTR --- 0 AODV 48 [0 4000000 ffff0008 0] ---- --- [4:255 -1:255 26 0] [0x2 5 1 [19 0] [0 4]] (REQUEST)
s 2.019681040 _5_ RTR --- 0 AODV 48 [0 4000000 ffff0008 0] ---- --- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
s 2.021656040 _5_ MAC --- 0 AODV 48 [0 5000000 ffff0008 0] --- ---- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
r 2.021848206 _4_ MAC --- 0 AODV 48 [0 5000000 ffff0008 0] --- ---- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
r 2.021848206 _6_ MAC --- 0 AODV 48 [0 5000000 ffff0008 0] --- ---- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
r 2.021873206 _4_ RTR --- 0 AODV 48 [0 5000000 ffff0008 0] ---- --- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
r 2.021873206 _6_ RTR --- 0 AODV 48 [0 5000000 ffff0008 0] ---- --- [5:255 -1:255 25 0] [0x2 6 1 [19 0] [0 4]] (REQUEST)
s 2.023514660 _6_ RTR --- 0 AODV 48 [0 5000000 ffff0008 0] ---- --- [6:255 -1:255 24 0] [0x2 7 1 [19 0] [0 4]] (REQUEST)
s 2.024139660 _6_ MAC --- 0 AODV 48 [0 6000000 ffff0008 0] --- ---- [6:255 -1:255 24 0] [0x2 7 1 [19 0] [0 4]] (REQUEST)

According to the detailed statistics records, RREQ_r number of node 3 is 1, RREQ_r number of node 4 is 1, RREQ_r number and RREQ_s number of node 5 and

node 6 are 1.Note that in the routing request packet number statistics, if only considering the routing layer data communication, the data records are identified as the fourth field in the "MAC" records do not conform to the requirements, identified as "RTR" on behalf of the routing layer data record is the record of this article needs. So, the RREQ_s number of node 5 is 1 instead of 2, by the same token, the RREQ_r number of node 4 and node 6 is 1 rather than 2.

In order to effectively detect the routing replay attack, the “Cross-layer” method is used to choose feature vector. According to different layer, the refinement of feature vector as shown in table 3. the MAC layer of CBR package delivery number, routing layer receives the number of CBR packages, the routing request packet routing layer number of RREQ and RREP response are chosen as feature vector of detecting routing replay attack, taking a unit time interval eigenvector statistics.

Table 3. routing replay attack feature format

feature	CBR_MAC_s	CBR_RTR_r	RREQ	RREP
format	1：n_1	2：n_2	3：n_3	4：n_4

Table 3 represents a period of time of a node is in the light of different packet types. In order to reduce the amount of calculation and improve the efficiency of detection, when choosing the characteristics of the raw data vector, not all feature vectors are counted, only some representative feature vectors are did for SVM training.

After the routing reply attack feature vectors ready, the awk program should be write according to the requirement. Detailed traversal raw data, each node can be got normal state and routing reply attack state eigenvector count.

Using awk programs on normal simulation data and simulation data of routing reply attack feature vector extraction, 2001 samples from positive and negative feature vector are got. As shown in figure 3, it is statistical feature vector data samples every 0.5 s under the routing replay attack from node 15.

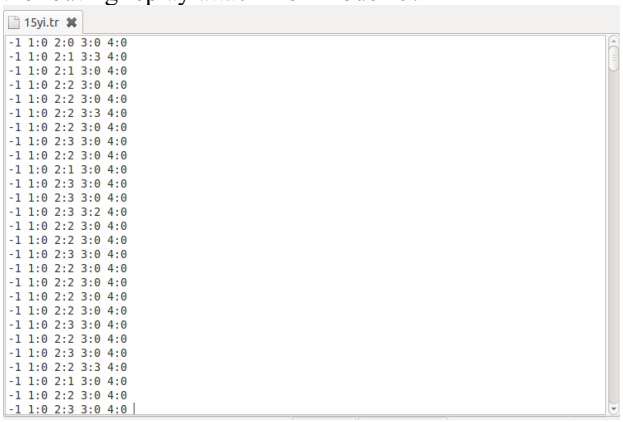

Figure 3. routing replay attack feature vector samples

3 THE SVM TRAINING PARAMETER OPTIMIZATION

Using libSVM software to training samples parameter optimization. First of all, the extracted feature vector samples can be divided into two parts, some are used for parameter optimization and the SVM training, the other part is used as the SVM classifier classification effect of a test. Training parameter optimization process and results shown as figure 4 and figure 5, the final optimal parameters: $C=0.5,\gamma=0.0078125,\sigma=128$, rate of cross validation is 99.95%.

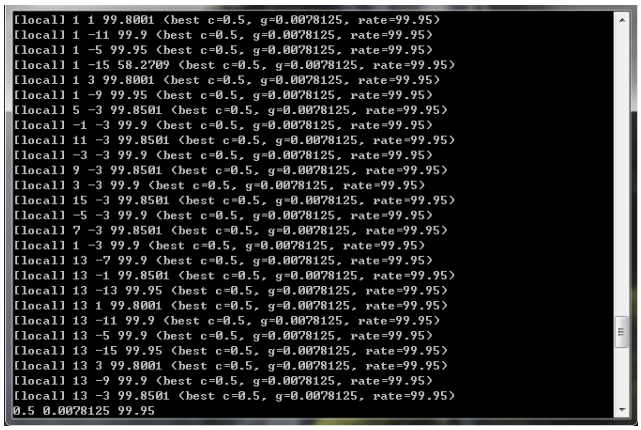

Figure 4. routing replay attack parameter optimization

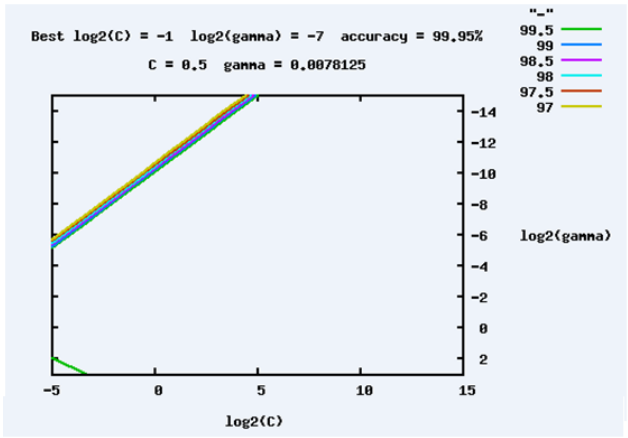

Figure 5. routing replay attack optimization result

4 Intrusion detection classifier training results and analysis

The optimal parameters and the positive and negative feature vector are input to SVM training, the routing replay attack intrusion detection classifier model as shown in figure 6.

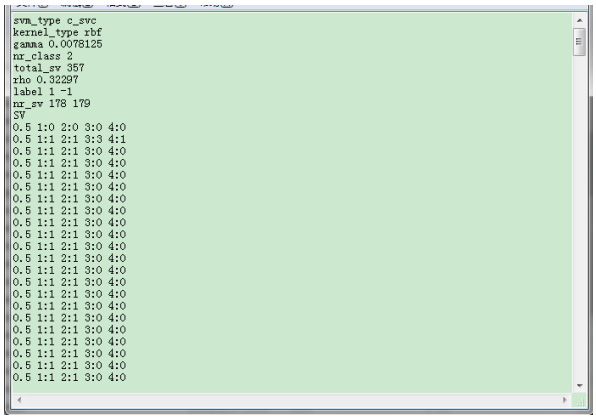

Figure 6. intrusion detection classifier model

Reoccupy has trained SVM classifier model forecast test sample, the prediction results are obtained as shown in figure 7.

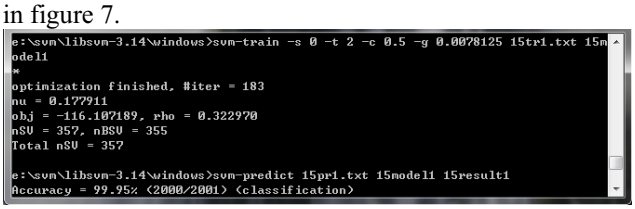

Figure 7. prediction results

From the result, Number of iterations of the intrusion detection classifier training algorithm is 183. The offset item in the discriminant function is 0.322970. The value of objective function is -116.107189. The number of support vector is 357, and the positive number is 178, the negative number is 179. The prediction accuracy reach 99.95%.

The simulation results of routing replay attack intrusion detection show that, on the choice according to the characteristics of the routing replay attack vector, the selected four features for routing replay attack detection is feasible, the establishment of VANET routing replay attack IDS based on SVM can obtain better testing result.

References

1. E.R. Gilles, B. Martine, P. Samuel, Q. Alejandro. Computer Communications, 44: 1-13 , 2014
2. G.J. Yan， O. Stephan, C.W. Michele. Computer Communications, 12: 2883-2897, 2008
3. T.W. Chim, S.M. Yiu. IEEE Transactions on Computers, 2: 510-524, 2014
4. J.A. Martinez, D. Vigueras, F.J. Ros, P.M. Ruiz. Communications and Networks, 2: 122-131, 2013
5. L.Y. Yeh, C.C. Yang, J.G. Chang, Y.L. Taai. Journal of Network and Computer Applications, 1: 284-292, 2013
6. M.M. Adankon, M. Cheriet. The Journal of the Pattern Recognition Society, 12: 3264-3270, 2009
7. G. Ratsch, S. Mika, B. Scholkopf, K.R. Muller. IEEE Transactions on Pattern Analysis and Machine Intelligence, 9: 1184-1199., 2002
8. W. Ju, H.D. Cheng. New Mathematics and Natural Computation, 1: 27, 2013
9. Y.C. Xiao, H.G. Wang, L. Zhang, W.L. Xu. Knowledge-Based Systems, 59: 75-84, 2014
10. R.S. Hamid, E. Mahdieh. Water Resources Management, 7: 2623-2644, 2013
11. A. Hamieh, O.J. Ben, L. Mokdad. Global Telecommunications Conference, 1-5, 2009
12. P.M. Nishant, K.V. Sunil, D.D. Jayesh. International Journal of Engineering Trends and Technology, 5: 2151-2157, 2013
13. D. Thierry, M. Nathalie, I. Sergio, H. Thomas. Mobile Information Systems ,4: 329-359, 2011
14. Y.G. Zhang, W.K. Lee, Y. Huang. Wireless Networks, 5: 545-556, 2003