INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

TRAFFIC PATTERN ANALYSIS IN MANET USING SEGMENTATION OF NETWORK

M.Vignesh[1], V. Parvathinathan[2]
[1]PG Scholar, [2]Assistant Professor,
SCAD College of Engineering and Technology, Tirunelveli, India

Abstract

Mobile Ad Hoc Network (MANET) is self-configuring network which can work without any centralized management. In MANET each node can perform both as host and router. Secure communication is entrusted by implementing Anonymous Routing Protocols which protect the identities of sources, destinations and Traffic pattern of the nodes from the attackers. But there are some other techniques were adapted for finding the source identities, destination identities and Traffic Pattern of the Network. A novel Traffic Pattern Analysis method is implemented to perform Traffic Analysis and the Source and Destination identities can be discovered. This system works based on the Source/ destination Probability distribution of the Packets. It works passively that is, the required task is performed without decrypting the packets and without interrupting the Network infrastructure.

1. Introduction

MANET is an infrastructure less, wireless and self-configuring network of mobile devices without any centralized management. These are mainly used in the military field. It is easy to deploy the network without any preplanned. In MANET to improve the secure communication, Anonymous Communication is used. Anonymous communication hides the link between the source and destination. It is difficult to find the source or destination of the communication link and the other intermediate nodes involved in it and finding the information or data flow through the network. For carrying out anonymous communication in the MANET many anonymous routing protocols are used in ad hoc routing such as MASK (Zhang et al. 2006), OLAR (Qin & Huang 2008) etc.

To improve the anonymity of the communication many techniques are used like onion routing (Reed et al.2002) with the MASK and OLAR protocols which includes the multiple layers of encryption of data. It hides the routing information and identity of nodes from the attacker nodes. The anonymity enhancing techniques are used to protect the MANETs. The routing information's are detected via the passive attacks which will not interrupt the network environment. Traffic analysis is used to track the data. There are many traffic analysis methods available but they are not well efficient in analyzing the traffic because of the three natures of MANET. They are broadcasting nature, ad hoc nature and mobile nature. By means of broadcasting nature, the packets are transmitted among nodes where, it is difficult to identify sender and receiver. By means of ad hoc nature, it is more possible for a node to behave both as sender and receiver. By means of mobile nature, it is predicted that the nodes are in mobility; which leads to the MANET environment to be more complex, for performing analysis.

2. Related Work

In the Traffic Inference Algorithm (TIA) for MANET, the difference between data frames, routing frames and MAC frames are exposed to the passive attackers. It allows the attackers to analyze the point to point traffic using MAC control frames, the end to end traffic using routing frames and the actual data or traffic pattern utilizing data frames. Traffic analysis in anonymous MANETs (He et al. 2008) and Traffic inference in anonymous MANETs (Liu et al. 2010) are two good approaches which based on deterministic network behaviors.

3. Proposed System

To analysis the hidden traffic pattern in MANET, Segmentation technique includes two steps 1. It constructs the point-to-point matrices and followed by constructing the end-to-end matrix from the packets. 2. Using the end-to-end matrix, it calculates the probability for each node which may be the source or destination (source/ destination probability) and it is considered that each pair of node take part in an end-to-end communication link (end-to-end link probability). That the two packet captured in different location in different time may be same packet to avoid this segmentation is implemented. In Figure 1.1 simple example is given. The Figure 1.1 depicts the model of simple mobile ad hoc network. There are four mobile nodes namely 1, 2, 3, 4. As per the figure the node 2 lies in the transmission range between node 1 and node 3, which performs the functions as receiving the packet from node 1 and forwarding the packet to node 3. The node 3 lies in the transmission range between node 2 and node 4, which performs the functions as receiving the packet from node 2 and forwarding the packet to node 4. If any of the nodes is not in the transmission range of any nodes then transmission between those nodes will not occur.

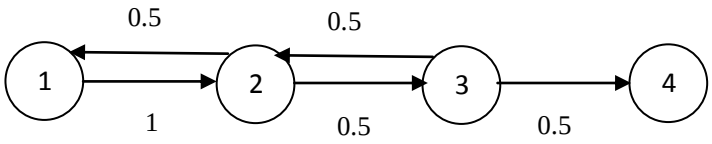

Figure 1.1: Simple Mobile Ad Hoc Network

It is considered that the node which has the probability of sending packet is 1. By this assumption node 1 sends packet to node 4 via node 2 and node 3. Hence node 1 has the probability 1 for sending packet to node 4 via node 2 and node 3. As the node 2 has probability 0.5 for sending packet to node 1 and probability 0.5 for sending packet to node 3. The node 3 has probability as 0.5 for sending packet to node 2 and probability 0.5 for sending packet to node 4.

3.1 Point-to-Point Matrices

The Point-to-Point Matrices are constructed between the neighbor nodes. P_1 Matrix is constructed between the node 1 and node 2. P_2 Matrix is constructed between the node 2 and node 3. P_3 Matrix is constructed between the node 3 and node 4.

$$P_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad P_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$P_3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

These matrices are constructed by the packet transmission between the nodes as node m to node n. For example in the P_1 Matrix 1st row vs. 1st column denotes the value set as per the probability given (0 or 0.5 or 1) packet transmission between node 1 to node 1 occurs and the 1st row vs. 2nd column denotes the value set as per the probability given (0 or 0.5 or 1) packet transmission between the node 1 to node 2 occurs. By using this procedure the Matrices are constructed.

3.2 End-to-End Matrix

The End-to-End Matrix is constructed from the Point-to-Point Matrices. All the Point-to-Point Matrices are combined and the End-to-End Matrix is constructed. In this End-to-End Matrix the packet transmission between node 1 to node 4 is calculated as $P_{1,4} = \min\{P_{1,2}, P_{2,3}, P_{3,4}\} = \min\{1, 0.5, 0.5\} = 0.5$ and the remaining node packet transmission is calculated as $P_{1,3} = \min\{P_{1,2}, P_{2,3}\} = \min\{1, 0.5\} = 0.5$, $P_{2,4} = \min\{P_{2,3}, P_{3,4}\} = \min\{0.5, 0.5\} = 0.5$, $P_{3,1} = \min\{P_{3,2}, P_{2,1}\} = \min\{0.5, 0.5\} = 0.5$, $P_{4,1} = \min\{P_{4,3}, P_{3,2}, P_{2,1}\} = \min\{0, 0.5, 0.5\} = 0$ and $P_{4,2} = \min\{P_{4,3}, P_{3,2}\} = \min\{0, 0.5\} = 0$.

$$E = \begin{bmatrix} 0 & 1 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Then the source probability distribution vector and destination probability distribution vector will be calculated for all nodes to be a source node or destination node. For calculating the source node probability distribution vector d_0 is taken as $(1/N, 1/N, 1/N, 1/N)$ here N denotes the number of nodes in the Network so $N = 4$, $d_0 = (1/4, 1/4, 1/4, 1/4)^T$.

$$s_v = \sum_{n=1}^{N} e(m,n) \times d_0 \qquad (3.1)$$

$$e(m,n) = \begin{bmatrix} 0 & 1 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$d_0 = (1/4, 1/4, 1/4, 1/4)^{\mathrm{T}}$$

From the equation source probability distribution vector will be calculated and all the probability distribution vectors in this paper are normalized. 'Normalizing' a vector means dividing each element of the vector by the summation of all the elements of the normalized vector equal to 1. The source probability distribution vector will be $s_v = (0.4, 0.3, 0.3, 0)^{\mathrm{T}}$. Then the destination vector will be calculated by using the following equations

$$d_v = \sum_{n=1}^{N} e(n,m) \times s_v \qquad (3.2)$$

$$e(n,m) = \begin{bmatrix} 0 & 0.5 & 0.5 & 0 \\ 1 & 0 & 0.5 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0.5 & 0 \end{bmatrix}$$

$$s_v = (0.4, 0.3, 0.3, 0)^{\mathrm{T}}$$

It is determined that the destination probability distribution vector from the above equation is $d_v = (0.18, 0.32, 0.21, 0.29)^{\mathrm{T}}$.

3.3 Source Node Probability

The Source node probability will be calculated from the End-to-End Matrix and source probability distribution vector.

$$S = (\Phi(E) \cdot \Phi^{\mathrm{T}}(E)) \cdot s_v \qquad (3.3)$$

Define a function $\Phi(E) = e(m,n) \times c(m,n)$, and then $\Phi^{\mathrm{T}}(E)$ denotes the transpose of $\Phi(E)$.

$$c(m,n) = c(n,m) = 1 - \frac{Sim\,(O\,(m), O(n) + Sim\,(I(m), I(n))}{2}$$

where O(m) and O(n) denote the m[th] row and n[th] row in E(the outgoing traffic from m and n), while I(m) and I(n) denote the m[th] and n[th] column in E(the incoming traffic to m and n). The vector space similarity (or cosine similarity) of two vectors V and U is defined as follows:

$$Sim(V,U) = V \cdot U / (|V||U|),$$

Where V · U denotes the dot product of V, and U, |V|, and |U| denote the norm of V and U.

$$c(m,n) = c(n,m) = 1 - \frac{Sim\,(1,0) + Sim\,(1,1.5)}{2}$$

$$c(m,n) = 0.7$$

$$\Phi(E) = e(m,n) \times c(m,n) = \begin{bmatrix} 0 & 1 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times (0.7)$$

$$S = (\Phi(E) \cdot \Phi^{\mathrm{T}}(E)) \cdot s_v$$

$$S = \begin{bmatrix} 0 & 0.7 & 0.35 & 0.35 \\ 0.35 & 0 & 0.35 & 0.35 \\ 0.35 & 0.35 & 0 & 0.35 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0.35 & 0.35 & 0 \\ 0.7 & 0 & 0.35 & 0 \\ 0.35 & 0.35 & 0 & 0 \\ 0.35 & 0.35 & 0.35 & 0 \end{bmatrix} \times \begin{bmatrix} 0.4 \\ 0.3 \\ 0.3 \\ 0 \end{bmatrix}$$

$$S = (0.44, 0.26, 0.30, 0)^{\mathrm{T}}$$

From the Source node probability, the node 1 has the highest probability compared with other nodes. So the node 1 is considered to be the most possible source node. It reveals that the node 4 has Zero probability because node 4 does not transmit any packet in the network.

3.4 Destination Node Probability

The Destination node probability will be calculated from the End-to-End Matrix and destination probability distribution vector.

$$D = (\Phi^{\mathrm{T}}(E) \cdot \Phi(E)) \cdot d_v \qquad (3.4)$$

Define a function $\Phi(E) = e(m,n) \times c(m,n)$, and then $\Phi^{\mathrm{T}}(E)$ denotes the transpose of $\Phi(E)$.

$$c(m,n) = c(n,m) = 1 - \frac{Sim(O(m), O(n) + Sim(I(m), I(n))}{2}$$

where O(m) and O(n) denote the m[th] row and n[th] row in E(the outgoing traffic from m and n), while I(m) and I(n) denote the m[th] and n[th] column in E(the incoming traffic to m and n). The vector space similarity (or cosine similarity) of two vectors V and U is defined as follows:

$$Sim(V,U) = V \cdot U / (|V||U|),$$

Where V · U denotes the dot product of V, and U, |V|, and |U| denote the norm of V and U.

$$c(m,n) = c(n,m) = 1 - \frac{Sim(1,0) + Sim(1,1.5)}{2}$$

$$c(m,n) = 0.7$$

$$\Phi(E) = e(m,n) \times c(m,n) = \begin{bmatrix} 0 & 1 & 0.5 & 0.5 \\ 0.5 & 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times (0.7)$$

$$D = (\Phi^{\mathrm{T}}(E) \cdot \Phi(E)) \cdot d_v$$

$$D = \begin{bmatrix} 0 & 0.35 & 0.35 & 0 \\ 0.7 & 0 & 0.35 & 0 \\ 0.35 & 0.35 & 0 & 0 \\ 0.35 & 0.35 & 0.35 & 0 \end{bmatrix} \times \begin{bmatrix} 0 & 0.7 & 0.35 & 0.35 \\ 0.35 & 0 & 0.35 & 0.35 \\ 0.35 & 0.35 & 0 & 0.35 \\ 0 & 0 & 0 & 0 \end{bmatrix} \times \begin{bmatrix} 0.18 \\ 0.32 \\ 0.21 \\ 0.29 \end{bmatrix}$$

$$D = (0.16, 0.34, 0.20, 0.29)^{T}$$

From the Destination node probability, the node 2 has the highest probability compared with other nodes. So the node 2 is considered to be the most possible Destination node.

4. Simulation Results

The network environment is simulated using Network Simulator 2(NS2). In this MANET environment 10 mobile nodes are randomly deployed in an 800 × 800 m^2 area. There are two types of node identification is considered which can perform as source node identification and destination node identification.

4.1 Source node identification

Table 4.1: Source Node Identification

S. No	Nodes	Probability
1.	0	0
2.	1	0
3.	2	0.2
4.	3	0
5.	4	0.4
6.	5	0
7.	6	0.2
8.	7	0.2
9.	8	0
10.	9	0

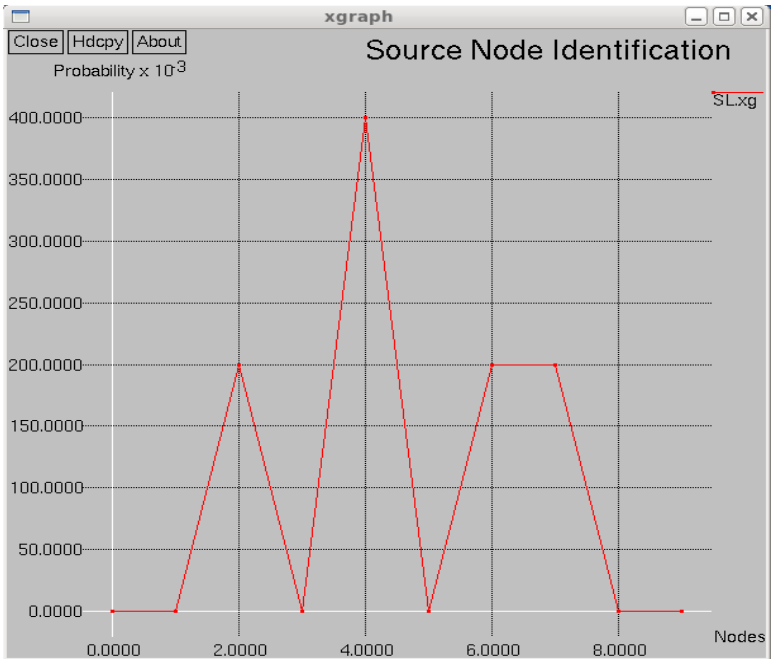

Figure 4.1: Source Node Identification

It depicts (Figure 4.1 & Table 4.1) very clearly that the node 4 secure the highest probability value (0.4) when compare with other nodes in the network. It is understand that the node 4 behaves as possible source node.

4.2 Destination node identification

Table 4.2: Destination Node Identification

S. No	Nodes	Probability
1.	0	0.1
2.	1	0.1
3.	2	0
4.	3	0.1
5.	4	0
6.	5	0.2
7.	6	0
8.	7	0
9.	8	0.4
10.	9	0.1

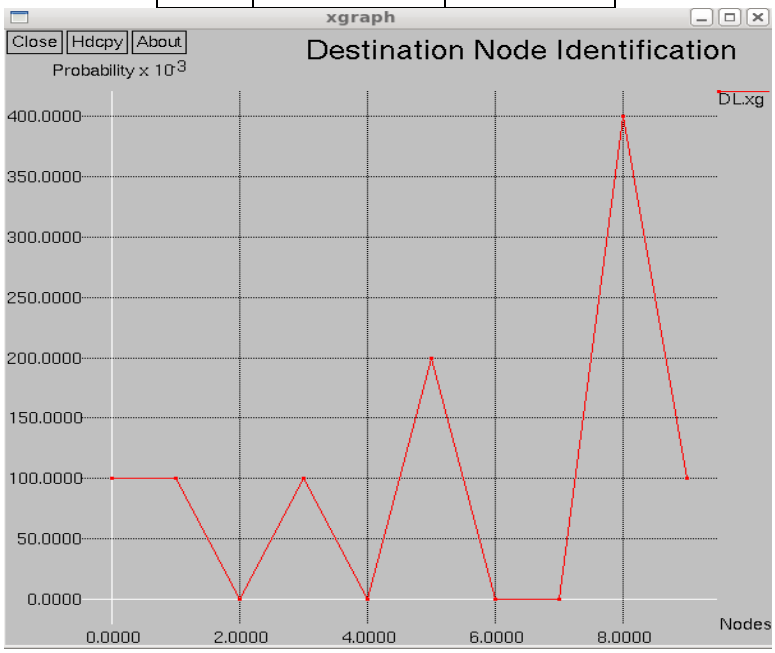

Figure 4.2: Destination Node Identification

It depicts (Figure 4.2 & Table 4.2) very clearly that the node 8 secure the highest probability value (0.4) when compare with other nodes in the network. It is understand that the node 8 behaves as possible destination node.

5. Conclusion

Traffic Pattern Analysis in MANET using segmentation is implemented for finding the source, destination identities and traffic pattern in MANET. This System constructs point-to-point matrices to derive end-to-end matrix from the probability of the packets. The source, destination identities and traffic patterns has been identified from the end-to-end matrix. The source, destination identities and traffic patterns are found by this system without decrypting the packets and without interrupting the network. The hidden traffic Pattern can be discovered ever without the knowledge about the actual source, destination and end-to-end communication relations.

REFERENCES

1. Blaze. M, Ioannidis. J, Keromytis. A, Malkin. T, and Rubin. A (2004). 'WAR: Wireless Anonymous Routing', Proc. Int'l Conf. Security Protocols, pp. 218-232.
2. Boukerche. A, El-Khatib. K, Xu. L, and Korba. L (2004). 'SDAR: A Secure Distributed Anonymous Routing Protocol for Wireless and Mobile Ad Hoc Networks', Proc. IEEE 29th Ann. Int'l Conf. Local Computer Networks , pp. 618-624.
3. He. T, Wong. H, and Lee. K (2008). 'Traffic Analysis in Anonymous MANETs', Proc. Military Comm. Conf. (MILCOM '08), pp. 1-7.
4. Huang. D (2008). 'Unlinkability Measure for IEEE 802.11 Based MANETs', IEEE Trans. Wireless Comm., vol. 7, no. 3, pp. 1025-1034.
5. Liu. Y, Zhang. R, Shi. J, and Zhang. Y (2010). 'Traffic Inference in Anonymous MANETs', Proc. IEEE Seventh Ann. Comm. Soc. Conf. Sensor Mesh and Ad Hoc Comm. and Networks (SECON '10), pp. 1-9.
6. Qin. Y and Huang. D (2008). 'OLAR: On-Demand Lightweight Anonymous Routing in MANETs', Proc. Fourth Int'l Conf. Mobile Computing and Ubiquitous Networking, pp. 72-79.
7. Reed M., Syverson and Goldschlag D. (2002), 'Anonymous Connections and Onion Routing', IEEE J. Selected Areas in Comm., vol. 16, no. 4, pp. 482-494.
8. Seys. S and Preneel. B (2006). 'ARM: Anonymous Routing Protocol for Mobile Ad Hoc Networks', Proc. IEEE 20th Int'l Conf. Advanced Information Networking and Applications Workshops, pp. 133-137.
9. Yang Qin, Dijiang Huang, *Senior Member, IEEE,* and Bing Li (2014) 'STARS: A Statistical Traffic Pattern Discovery System for MANETs'.
10. Zhang. Y, Liu. W, Lou. W, and Fang. Y (2006). 'MASK: Anonymous On-Demand Routing in Mobile Ad Hoc Networks', IEEE Trans. Wireless Comm., vol. 5, no. 9, pp. 2376-2385.