International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169
Volume: 4 Issue: 5 391 – 395

Collaborative Approch and Design of Traffic Pattern Discovery System in Mobile Ad Hoc Network

Ayesha Khan
Department of Computer Science and Engineering
Jhulelal Institute of Technology, Nagpur
Nagpur, Maharashtra, India
angel93255@gmail.com

Prof. Ajay Karare
Department of Computer Science and Engineering
Jhulelal Institute of Technology, Nagpur
Nagpur, Maharashtra, India
ajju.karare@gmail.com

***Abstract*—**Mobile Ad-hoc Network (MANET) is a collection of loosely connected nodes through wireless links. The mobile nature and dynamic route formation of the network exposes it to variety of attacks. MANET finds its application in wide range of domains which includes military environment where confidentiality and detecting the activities of groups indulged in war crimes. This includes detecting the identity of anonymous nodes in the network. Traffic analysis type of passive attack on MANET aims to discover the identity of such anonymous nodes. Traffic Analysis attack monitors the packets communicated between the nodes and works fine if every packets are encrypted. The proposed system targets to disclose the identity of anonymous nodes in a MANET where Best- First Search is employed for route formation. The proposed work will find out the possibilities of a node being a sender or receiver for the analyzed traffic. The results can be obtained in matrix form where each row will refer to each node. The work completed up till now includes analysis of the traffic in static manner and finding out the source and destination of the packets transmitted. Network Simulator-2 (NS2) is used for simulating the results.

Keywords-*Anonymous node;Best-First Search; mobile ad-hoc network (MANET);network simulator-2 (NS2); traffic analysis*

*****

I. INTRODUCTION

Mobile Ad-hoc Network (MANET) is referred to as self-configuring, self-organizing and infrastructure less network. Fig. 1 illustrates the basic structure of MANET. These properties allow MANET to be widely used in military environment. Security is the prime concern in MANET, especially in military application. Due to lack of any central coordination and shared wireless medium makes MANET vulnerable to attacks. There are two types of attacks: Active Attack and Passive Attack.

Traffic analysis [13] is one of the types of passive attacks. This attack is required in military environment. The main motives of soldiers are to detect the communication path of the adversarieswithout their knowledge. Predecessorattack [11] anddisclosure attack [12] are the types of traffic analysis attack. One of the major issues of MANET is communication anonymity. An anonymous communication system can be defined as a technology that hides the object identity. There are two aspects of communication anonymity: Source/Destination Anonymity and End-to-End Anonymity [1, 14].Source/Destination Anonymity: There is difficulty in identification of source and destination nodes in the network.End-to-End Anonymity: There is difficulty in identification of end-to-end communication relation.

Earlier, communication anonymity has been proposed by anonymous routing protocols. These include Anonymous On-Demand Routing (ANODR) [10] and On-demand Lightweight Anonymous Routing (OLAR) [9]. These anonymous routing protocols hide the information by encrypting the packet from attackers. However, predecessor and disclosure attacks were used to intercept the information. The following are the nature of MANET due to which traffic analysis becomes difficult task:1) Broadcasting nature: Point-to-Point transmission is applied to only one receiver in wired network. While in wireless network message is broadcasted to multiple receivers. 2) Ad-hoc nature: There is confusion in the role of a mobile node as it can serve as source as well as destination.3) Mobile nature: Traffic analysis model do not consider the mobility of communication peers. This makes the communication among mobile nodes more complex.

In order to curb above characteristics there is requirement of a technology. The primary objective of this proposed system is to show that statistical traffic analysis can be performed without the knowledge of adversary.

The remaining paper is organized as follows: Section II describes the related work. Section III presents the proposed system. Section IV describes the implementation of the proposed system. Last section i.e.section V presents the conclusion.

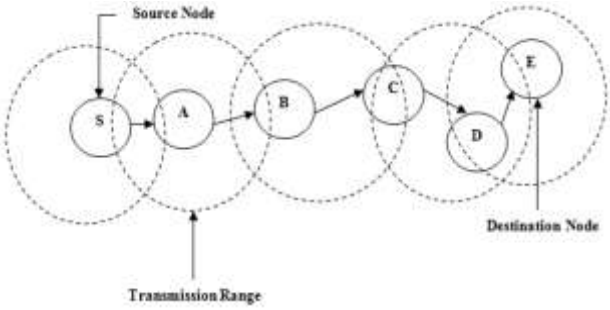

Fig. 1 Basic Structure of Mobile Ad hoc Network

II. RELATED WORK

Qin et al. [1], proposed that even though there are many anonymous routing protocols and anonymous enhancing techniques available, mobile ad hoc network (MANET) is vulnerable to passive statistical traffic analysis attacks The primary objective of STAR (Statistical Traffic Pattern Discovery System) is used to discover the hidden traffic pattern in MANET. However, this proposed system has not

included any searching algorithm that can be applied to search the traffic free path.

Lu and Liu [2], proposed Best-First Search is better alternative than depth first search (DFS) for Prolog engine. This system associate a heuristic function established on an extended measurement of UCB1. UCB1 (Upper Confidence Bound) is one of the machine learning algorithm used in Markovian Decision Process (MDP). This heuristic function can equalize exploitation and exploration during the search process carried out in Prolog engine. The following are the advantages of Best-First Search over depth first search (DFS) discussed by Benjie Lu et.al:

(i) In DFS if the solutions are nearer to search tree then they can be found efficiently otherwise it will take a long route to reach the desired path. This will affect the time required for execution.

(ii) Sometimes solutions in DFS is not guaranteed as it may get trapped in an endless branch of a search tree. There is no such problem in best-first search.

(iii) All the above discussed problems will affect the efficiency of DFS. Thus efficiency of best-first search is comparatively better than DFS.

From TABLE I we conclude that Best-First Search is better searching algorithm than DFS.

Best-First Search is relevant to MANET due to following advantages: 1) It takes less time to generate solution even if problem space is large. 2) When one path is selected others are kept around so that they can be revisited later if the selected path becomes less promising. 3) It tries to find out path with lowest cost. 4) It can be an option for complex problems.

Kelly et al. [3], thoroughly investigated on anonymity. Anonymity for a user can be defined as using any services while keeping their identity hidden from an adversary.The user can protect their data from attacks with the help of anonymity.The three properties of anonymity areUnidentifiability, Unlinkability, and Unobservability.Unidentifiability is subdivided into sender anonymity (SA), receiver anonymity (RA), mutual anonymity (MA) andgroup anonymity (GA).

TABLE I. COMPARISON OF SEARCHING ALGORITHMS

Parameters	**Searching Algorithms**	
	Best-First Search [2]	***Depth First Search [2]***
Time required for searching	Less	More
Solution	Guaranteed	Not guaranteed
Efficiency	More	Less

Our primary objective will be to discover the traffic pattern. So, we have to work on unidentifiability property of anonymity and decrease the sender anonymity (SA) and receiver anonymity (RA).

Liu et al. [5], proposed that traffic intensity and packet number are the two important metrics in order to detect the attack in a network. The proposed anomaly detection system is used to detect the distributed denial of service (DDoS) attack in MANET. These two metrics are used to detect the DDoS attack whenever traffic analysis is carried on MANET. Similarly when traffic analysis will be carried on our proposed system required parameter will be data transmission.

Zhang and Zhang [6], recommended that for route discovery in MANET, control traffic plays an important role. When on demand routing protocols are used the characteristic that involve to carry out research on control traffic in MANET include distribution of nodes' control packet traffic, communication of control packets between nodes, rate of RREQ (route request) packets and the ratio of number of RREQ packets originating from one node to all RREQ packets relayed by this node. The factors such as mobility, node density and data traffic affects these characteristics of control traffic. Thus theoretically we can conclude that one of the required factors of control traffic is data traffic. Hence in our proposed system control traffic will also play an important role in route discovery

Liu et al. [7], designed traffic inference algorithm (TIA) which allows an adversary to interpret the traffic pattern in MANET. In this algorithm it is assumed that the difference between data frames, routing frames and MAC control frames is visible to passive adversaries. From these differences an adversary can identify the point-to-point traffic using the MAC control frames, recognize the end-to-end traffic by tracing the routing frames and then find out the actual traffic pattern using the data frames. The drawback of this algorithm is that it depends on the deterministic network behaviors.

Tehraniand Shahnasser [8], discussed on anonymity and importance of using anonymous communication protocol. They researched on best anonymous routing protocol and evaluated pros and cons of current solution, shortcoming and challenges that are not provided by these protocols.

Sen and Sahare [14], reviewed on discovery of traffic pattern in MANET. They are using a heuristic approach to find the hidden nodes. Earlier they studied on depth first search (DFS) for route discovery. However, best-first search is better than DFS is proved in TABLE I.

III. PROPOSED SYSTEM

The main aim of the proposed system will be to unhide the identity of the nodes. This will be implemented by using statistical traffic analysis. A heuristic searching algorithm i.e. Best-First Search will be used to traverse the searching path in the network. Then statistical traffic analysis will be performed on searched nodes. This will provide an estimation of datatransmitted to all the neighboring nodes of every searched node. Probability distribution will be used to discover the traffic pattern. The working of each module is explained below:

A. *Topology Formation and Setup Phase*

This step will mainly focus on the packet distribution between the nodes in MANET. Every node will send topology discover packet to connect with the neighboring nodes. The function of topology discovery packet is to inform a particular node's presence to all its neighboring nodes. The main focus will be on simulating the characteristics of MANET like, the mobile nature of the nodes. It will aim to setup connection in all the nodes available in the network.

B. *Route discovery using Best-First Search*

In military environment to get the communication path of adversary's, time is one of the important factors to be considered. Uninformed search algorithms cannot be used in

 ISSN: 2321-8169
391 – 395

such situation as they require too much time or space. Thus it is advisable to use heuristic search algorithms. The algorithm that uses heuristic functions for searching a particular node iscalled as heuristicsearchalgorithm. Heuristic search algorithms are efficient because they take advantage of feedback from the data to direct the search path. There are six heuristic search algorithms: Generate-and-Test, Hill Climbing, Best-First Search, Problem Reduction, Constraint Satisfaction and Means-Ends Analysis. The TABLE II shows the reason to choose best-first search algorithm from all the heuristic search algorithms.From TABLE II we can conclude that best-first search will be the correct algorithm to select for routing.

Best-first search is a heuristic search algorithm that traverses a graph in search of one or more than one goal nodes. This algorithm combines the advantages of depth first search (DFS) and breadth first search (BFS). Thus by combining these two advantages best-first search finds asolution in less interval of time.In best-first search process, the most promising nodes will be selected from generated nodes.This will be done by applying heuristic function to each of the nodes.Heuristic function estimates least costpath fromnode n to goal node. The chosen node will be expanded by using the rules to generate its successors.

TABLEII. REASON TO CHOOSE BEST-FIRST SEARCH

Heuristic Search Algorithms	Reasons
Generate-and-Test	Takes long time to generate solution if problem space is large.
Hill climbing	Once a move is selected other moves are rejected never to be reconsidered.
Problem Reduction	The desired path may not be the one with lowest cost.
Constraint Satisfaction	Used only to solve crypt-arithmetic problems.
Means-Ends Analysis	Cannot be used to solve complex problems.

If one of them is solution then quit the process. If solution is not found then all the generated nodes will be added to set of nodes. The above process will be continued until desired node will be found.

C. *Traffic Analysis in network*

The traffic analysis of the network aims to identify the state of the network at each point of time. The technique that is used to find out the state of network after regular intervals of time is time slicing technique. The time slicing mechanism captures the snapshot of the network after each quantum time giving a n × n matrix where n is the number of nodes in the network.This traffic matrix will consists of traffic volume (number of packets) from one node to another. For example,

$$M = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Here 1 indicates that there is transmission of data (traffic volume) from node 1 to node 2 whereas 0 indicate that there is no transmission of data between the two nodes. The traffic matrix obtained from the above helps to deduce point-to-point and end-to-end traffic volume between each pair of nodes. Such a matrix will give one to one hop packets and two or more hop traffic captured in all the matrices between all the nodes in the network.

With the completion of traffic analysis of the network, aim of the system is to discover the actual source/destination probabilities i.e. to find actual source and destination nodes in the network. For computation of probability distribution following summation will be used which will give the probability of the selected node being a source/destination. It is required to consider the identical probability distributionmatrices should be initialized to S0=D0= (1/N, 1/N, 1/N…). The reason for employing the equal probability is that with no traffic taking place in the network all the nodes will have same chances of being a source and being a destination.

The equation required for source probability distribution is

$$s'(i) = \sum_{j=1}^{N} r(i,j) \times d_0(j), \qquad (1)$$

The equation required for destination probability distribution is

$$d_1(i) = \sum_{j=1}^{N} r(j,i) \times s'(j), \qquad (2)$$

Where s (i) is probability of node i being source, N is number of nodes in the network, r (i, j) is the accumulative traffic volume from node i to node j, $d_0(i)$ is the destination vector.

Fig. 2, illustrates the flow of the proposed system. When the system will start it will form a network. This network will consist of certain number of nodes. All the nodes will be

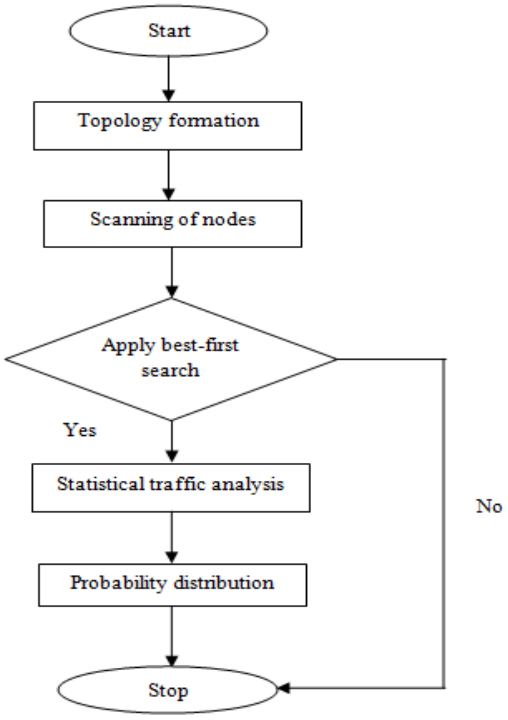

Fig. 2 Flow Diagram of Proposed System

scanned. In order to search the node a best-first search algorithm will be applied. If the required node is present then statistical traffic analysis will be performed on it. After performing statistical traffic analysis probability distribution will be applied to discover the traffic pattern. However, if the

required node is not found then the system will stop and no further process will be carried out.

D. *Data Routing*

After calculating point-to-point traffic matrix, end-to-end traffic matrix and probability distribution the next step will be to perform graphical analysis.The parameters included in graphical analysis will be throughput, delay, packet delivery ratio (PDR) and probability distribution. The graphical analysis will give a clear idea of analysis of each node according to above parameters. The most important graph will be probability distribution as it will give an analysis of a node being source or destination.

TABLE III. SIMULATION PARAMETERS

Parameter	**Value**
Routing Protocol	DSR
Simulation Time	30 seconds
Simulation Area	800x800 m^2
Number of Nodes	50
Traffic Type	UDP
Pause Time	0.5 seconds
Mobility	10 meter/sec
Packet Size	512 bits
Data Rate	512 kbps
Mobility Model	Random waypoint
MAC	802.11(a)
Channel Type	Wireless channel

IV. IMPLEMENTATION

A. *Simulation Environment*

NS2 is the simulation environment that is used to analyze the proposed system. Network Simulator-2 (NS2) is an event-driven simulation tool that is used to study the dynamic nature of communication network [4]. NS2 uses two languages: C++ as backend and Object-oriented Tool Command (OTcL) as frontend. There are tools such as NAM (Network AniMator) and X-Graph to interpret these results graphically and interactively [4]. The user can extract text-based data to evaluate a particular behavior of the network. The parameters used in this proposed system during the simulation process are given in TABLE III.

B. *Outcomes*

The following outputs are obtained during topology formation and setup phase:

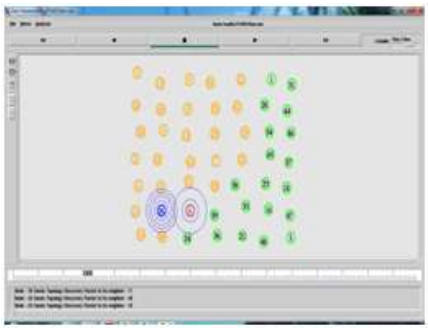

Fig. 3 Identification of topology

In Fig. 3, each node sends topology discovery packets to its neighboring nodes. This process is required to build connection with each of the neighboring nodes present in the network. During this phase the sender node turns into blue colour and the receiver node turns into brown colour. Brown colour also indicates that the node is in the process of receiving hello packets. Fig. 3, displays the data transmission between two nodes where sender node is node number 26 and receiver node is node number 6. Both the nodes maintain its colour until the process of sending hello packets is completed. Remaining nodes are displayed with orange colour which indicates that the nodes had completed topology formation phase.

Here in Fig. 4, we can see the transmission of data between nodes. The nodes that form path for data transmission are indicated with blue colour. Source node is node number 4 indicated by using blue colour. The intermediate nodes are labeled as N_1, N_2, N_3, N_4, N_5, and N_6 where N stands for node.N_1 is node number 9, N_2 is node number 30, N_3 is node number 25, N_4 is node number 20, N_5 is node number 7 and N_6 is node number 22. Destination node is node number 23 which is indicated with orange colour.

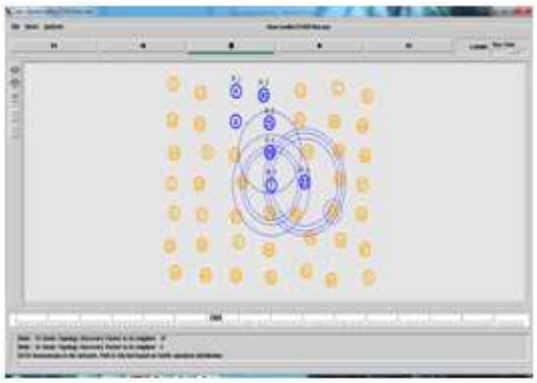

Fig. 4 Data transmission between nodes

Omni directional antennas are applied to send signals to the nodes. The signal flow is indicated with blue colour. Nodes are selected using Dynamic Source Routing (DSR) protocol.

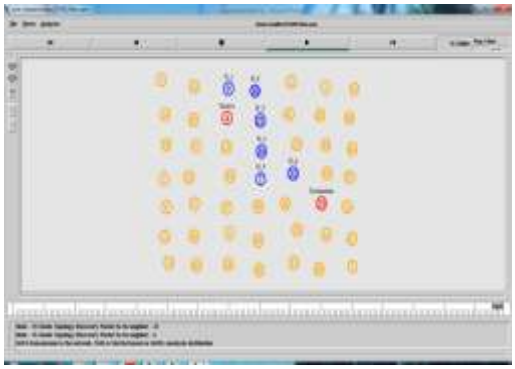

Fig. 5 Source and Destination nodes

The nodes in Fig. 5, are selected on the basis of traffic free node. Here we can observe a path from source node to destination node. The intermediate nodes are indicated withblue colour. The intermediate nodes are node number 9, 30, 25, 20, 7 and 22. They are labeled as N_1, N_2, N_3, N_4, N_5 and N_6. The source and destination nodes are indicated with red colour. Source node is node number 4 and destination node is node number 23. Node 4 and 23 are labeled as “Source” and “Destination” using black colour.

C. *Comparison of Performance Metrics on various Statistical Traffic Analysis*

The following TABLE IV shows theoretical comparison of performance parameters.

V. CONCLUSION

The main purpose of the proposed system is to disclose the identity of source and destination nodes in the network. This can be fulfilled by discovery of communication pattern. The communication pattern is discovered without decrypting the packets. This has been satisfied by using best-first search (a heuristic approach) for traversing the path, statistical traffic analysis for analyzing and to identify point-to-point transmission among receives. This is followed by calculating probability distribution to find out approximate source and destination nodes in the traced path. This reduces anonymous communication which is one of the characteristics in mobile ad hoc network (MANET).

The graphical results that are obtained from comparative analysis between all the three total numbers of nodes i.e. 25, 45 and 65 are discussed as follows: According to Fig. 12, the highest probability value for source obtained in 25 total numbers of nodes is 2 which are present between node numbers 10 to 15. For 65 total numbers of nodes the value is 6 which are present between node numbers 10 to 30. While, for 65 total numbers of nodes the value is 8 which are present between node numbers 20 to 40. For Fig. 13, the highest probability value for destination obtained in 25 total numbers of nodes is 4 which are present between node numbers 10 to 20. For 65 total numbers of nodes the value is 4 which are present between node numbers 25 to 35. While, for 65 total numbers of nodes the value is 7.5 which are present between node numbers 35 to 45. For Fig. 14, the overall success rate for different total number of nodes are 90 to 95% for 25, 100% for 45 and 90% for 65 total number of nodes. Similarly for Fig. 15, the overall success rates for different traffic pattern are 35% for 25, 100% for both 45 and 65 total numbers of nodes.

TABLE IV. COMPARISON OF PERFORMANCE METRICS

Parameters	**Types of Statistical Traffic Analysis**	
	Traffic Inference Algorithm	***Proposed System***
Packet Delivery Ratio (PDR)	Less	More
End to End Delay	More	Less
Throughput	Less	More

consideration various network parameters.

REFERENCES

[1] Yang Qin, Dijiang Huang and Bing Li "STARS: A Statistical Traffic Pattern Discovery System for MANETs" IEEE Transactions on Dependable and Secure Computing, Vol. 11, No. 2, March/April 2014.

[2] Benjie Lu and Zhingqing Liu,"Prolog with Best-First Search", IEEE 25th Chinese Control and Descision Conference, 2013.

[3] Douglas Kelly, Richard Raines, Rusty Baldwin, Michael Grimaila, and Barry Mullins, "Exploring Extant and Emerging Issues in Anonymous Networks: A Taxonomy and Survey ofProtocols and Metrics", IEEE Communications Surveys & Tutorials, Vol. 14, No. 2, Second Quarter 2012.

[4] Teerawat Issariyakul and Ekram Hossain, Introduction to Network Simulator NS2, Second Edition, Springer, 2012.

[5] Lei Liu, Xiaolong Jin, Geyong Min, and Li Xu, "Real-Time Diagnosis of Network Anomaly based on Statistical Traffic Analysis", IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications, 2012.

[6] Zhilin Zhang and Yu Zhang, "Control Traffic Analysis of On-Demand Routing Protocol in Mobile Ad-hoc Networks", IEEE Second International Conference on Networking and Distributed Computing, 2011

[7] Y. Liu, R. Zhang, J. Shi, and Y. Zhang, "Traffic Inference in Anonymous MANETs," Proc. IEEE Seventh Ann. Comm. Soc. Conf. Sensor Mesh and Ad Hoc Comm. and Networks, pp. 1-9, 2010.

[8] Tehrani.A.H. and Shahnasser. H.," Anonymous communication in MANET's, solutions and challenges," IEEE International Conference on Wireless Information Technology and Systems (ICWITS), pp. 1-4, 2010.

[9] Y. Qin and D. Huang, "OLAR: On-Demand Lightweight Anonymous Routing in MANETs," Proc.Fourth Int'l Conf. Mobile Computing and Ubiquitous Networking, pp. 72-79, 2008.

[10] J. Kong, X. Hong, and M. Gerla, "An Identity-Free and On-Demand Routing Scheme against Anonymity Threats in Mobile Ad Hoc Networks," IEEE Trans. Mobile Computing, vol. 6, no. 8, pp.888-902, Aug. 2007.

[11] M. Wright, M. Adler, B. Levine, and C. Shields, "The Predecessor Attack: An Analysis of a Threat to Anonymous Communications Systems," ACM Trans. Information and System Security, vol. 7, no. 4, pp. 489-522, 2004.

[12] G. Danezis, "Statistical Disclosure Attacks: Traffic Confirmation in Open Environments," Proc. Security and Privacy in the Age of Uncertainty, vol. 122, pp. 421-426, 2003.

[13] J. Raymond, "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems," Proc. Int'l Workshop Designing Privacy Enhancing Technologies: Design Issues in Anonymity Unobservability, pp. 10-29, 2001.

[14] Priyanka Sen and Vaishali Sahare,"A Survey on Traffic Pattern Discovery in Mobile Ad hoc Network," International Journal of Computer Science and Network (IJCSN) Vol.4, No. 1, pp. 25-30, 2015.