ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

IJARMATE
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

APPROXIMATE STATISTICAL TRAFFIC PATTERN DISCOVERY METHOD ON MOBILE ADHOC NETWORKS

Dr.B.Srinivasan[1], N.Prakash[2]
[1]Associate Professor of Computer Science, [2]Assistant Professor of Computer Science
Gobi Arts & Science College,
Gobi, Erode
0312prakash@gmail.com

Abstract—Data collection is the critical task in wireless sensor network applications, wireless sensor networks the approximate data collection process is the wise choice in communication bandwidth and energy budget. This paper focuses on efficient approximate data collection with specified errors in wireless sensor networks. The key idea of the data collection approach is Approximate Data Collection to divide a sensor network into clusters, discover local data correlations on each cluster head, and perform global approximate data collection according to model parameters uploaded by cluster heads. In the process of mode based data collection and formulate the problem of selecting the minimum subset of sensor nodes into a minimum dominating set problem which is known to be NP-hard, and propose a greedy heuristic algorithm to find an approximate solution. In addition, to provide periodic query scheduling for data aggregations with minimum delay under various wireless interference models.

Keywords: Ad-hoc Networks, Demand Routing, MANET, Traffic Pattern Discovery, Wireless Sensor Networks.

I. INTRODUCTION

In service-based systems designed for and deployed in MANETs, these data change frequently, as the system attempts to adapt to the dynamics of the underlying network bindings between services can dynamically change according to the status of the nodes, links, and paths. Typically, time-varying, local data are observed by *monitors* situated at the nodes of the network and captured as a *time series*. The time-series data are *harvested* from the nodes and presented the management element located in the network can perform a global state analysis, such as fault identification [16], [20], load re-balancing, or service re-placement.

The experiment examines two different mobility behaviours under the same hypothetical service based system. First behaviour is characterized by random, independent movements of nodes and the other is characterized by collective, grouped movements. The experiment models a management element, residing at some node in the network, needing access to the monitoring data stored on a specific subset of nodes in the network.

Privacy and security are emerged as an important research issue in mobile Ad-hoc Networks (MANET). The proposed Statistical traffic pattern discovery system (STARS) was introduced to discover the broadcast channels without altering the packet content as plaintext.

As mobile ad hoc network (MANET) systems are built to study MANETs has focused on developing new applications such as collaborative games, collaborative computing, messaging systems, distributed security schemes, MANET middleware, peer-to-peer file sharing systems, voting systems, resource management and discovery, vehicular computing and collaborative education systems. The increasing sophistication of end-user devices has led to an increase in the richness and complexity of the systems deployed on structure as interconnected and interdependent services, so called service-based systems.

As time progresses the nodes in that subset increasingly become unreachable meaning that less and less monitoring data are available to the management element. Interestingly, we can see that random node movements can cause more problems than grouped node movements. Intrusion is any set of actions that attempt to compromise the integrity, confidentiality, or availability of a resource system for the detection of such intrusions. Since the prevention techniques cannot be sufficient and new intrusions continually emerge, IDS is an indispensable part of a security system. IDS are introduced to detect possible violations of a security policy by monitoring system activities and responding to those intrusive. If detect the attack once it comes into the network, a response can be initiated to prevent or minimize the damage of the system and also helps prevention techniques improve by the information about intrusion techniques.

A. *Our Contributions*

The proposed approach uses evolutionary computation techniques to explore the MANET complex design space. Using the Genetic Programming (GP) technique to detect known attacks against MANETs and evaluated on simulated networks with varying mobility and traffic patterns. The GP effectively detects known attacks, flooding and route disruption attacks. Intrusion detection system (IDS) on MANETs suited to resource-constrained environments. That employs Multi Objective Evolutionary Computation (MOEC) techniques in order to discover trade-offs between non-functional and functional

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

properties, and optimize the objectives simultaneously during evolution. The main contribution is to evolve a set of programs for each attack offering different trade-offs between intrusion detection ability of evolved programs energy usage and also to investigate or to evolve separate programs for each attacks.

To improve the performance of a typical MANET, various routing protocols have been proposed by many network researchers. For conciseness only a selected set of literature that is indicative of the range of approaches used for improving and analyzing MANET routing performance is reported.

II. EVALUATION METHODOLOGY

The evaluation is to understand the performance of the harvesting method under various conditions. In particular, we are interested in how well the method can improve the reachability of monitoring data in the face of network dynamics. In this section we describe the method we use to evaluate our approach. The method is based on a case study in which we experiment with the problem of discovering service dependencies. After describing the case study, we detail the tools, metrics and scenarios used to conduct the experiments.

A. Harvesting service dependent data

For discovering dynamic dependencies among the distributed service components of MANET hosted applications [16]. The technique suffered from the problem that it assumes all nodes of interest are reachable, on demand, from the node where the dependence analysis is to be carried out. In fact, it is a common occurrence that not all nodes are reachable, which significantly reduced the effectiveness of the technique and inspired the design of our new harvesting method. In this case study, we evaluate how well the new method can improve the availability of the monitoring data and, thereby, the effectiveness of the dependence discovery technique.

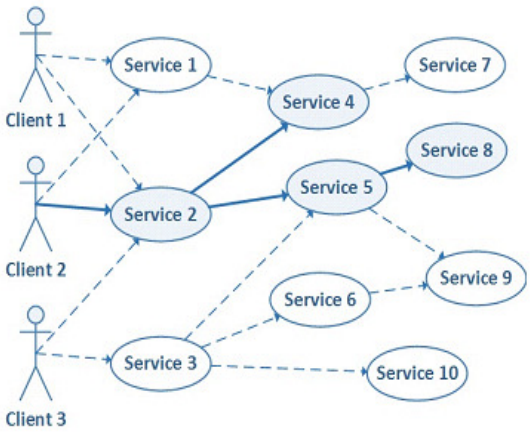

Fig.1. Evaluation structure of service model

Each time series represents the time-varying dependencies between a source and target, in which entries for each time slot are Boolean data about whether or not the given dependence occurred within that time slot. When the monitor detects the occurrence of dependence, it signifies this by setting a 1-bit flag in the corresponding time slot. It also records identifying information about the source and target of the dependence. The set of time slots thus represents an aggregated time series of dependencies. The set of relevant time slots shifts as new time slots are added and obsolete ones removed, reflecting the changes in dependencies.

B. Experimental setup

The application running within the MANET is built as a generic Web service system based on the Java Web Services stack Glassfish Metro.5. The system is composed of two kinds of configurable components, a generic client application and a generic Web service, structured as a 2-tiered system. The first tier consists of client-facing services, the second tier consists of backend services. The experimental uses 500 clients, fifty front-end services, and two hundred back-end services. Conversation of each client invokes a method selected uniformly at random from all methods are provided by the fifty front-end services. Each virtual CORE node runs a client /service, along with a monitor, a synchronization agent, and an analysis element. The analysis element would not normally be deployed to all nodes, but to give us maximum flexibility in evaluating the dissemination of monitoring data. The DGs are constructed on demand by the discovery element. The graphs are rooted at a given client, beginning at a given time instant, and for some time window. Each DG is constructed for a particular conversation time window begins and ends with the start and end of the conversation.

To collect our results from 40 minutes of execution after excluding 10 minutes of warm up. Each combination of parameters results in thousands of conversations during the 40-minute execution. The results given in the results section are averages over the data collected from these conversations. The primary evaluation question for our harvesting method depend the data of each conversation has been propagated through the network after a certain period of time.

III. SYSTEM ARCHITECTURE

The architecture of the proposed work consists of seven major components, namely, data collection agent, data Preprocessing module, intrusion detection system module and prevention module, user interface, decision manager, and knowledge base as shown in Figure 2.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

1) Data Collection Agent - collects the network data from the network layer and sent to the preprocessing module for preprocessing the data.

2) Data Preprocessing Agent - uses a preprocessing technique called attribute selection algorithm for effective preprocessing. In this technique, the agent selects only the valuable attributes from the data set using projection. Moreover, data cleaning, data integration, and data transformation are carried out for performing effective preprocessing.

3) Intrusion Detection Module - detects the intruders from the given data using intelligent agent-based weighted distance detection algorithm and intelligent agent-based enhanced multiclass support vector machine algorithm. The intrusion detection module distinguishes the intruders from normal users using an outlier detection algorithm combined with SVM for obtaining better classification accuracy.

4) Outlier Detection Agent - uses a newly proposed weighted-distance-based outlier detection algorithm of the agent uses an outlier factor to determine the outlier of a point in the feature set.

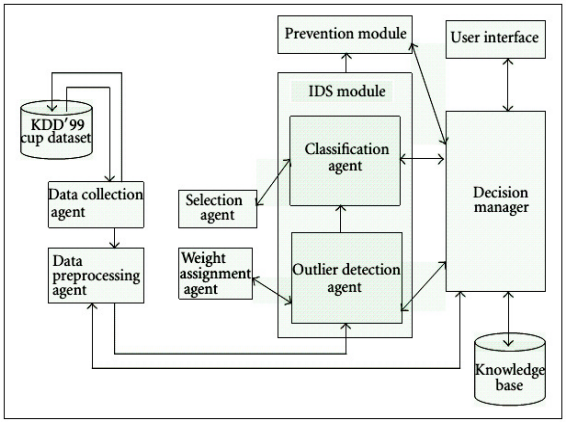

Fig.2. Frame work for system architecture

By privacy properties anonymity and resistance to track security properties include node / origin authentication and location integrity. The advantages of MANET are a strong privacy and strong security and easily create fraudulent phantom node location entries and propagate to the entire MANET. It is flexible and efficient but its drawbacks are do not consider jamming and denial-of- service (DoS) attacks. Such attacks are impossible to combat at the network layer.

B. Traffic Inference in Anonymous Manets

The open wireless medium in a mobile ad-hoc network (MANET) enables malicious traffic analysis to dynamically infer the network traffic pattern in hostile environments. The disclosure of the traffic pattern and its changes is often devastating in a mission-critical MANET. A number of anonymous routing protocols have been recently proposed as an effective countermeasure against traffic analysis in MANETs. The author propose a novel traffic inference algorithm, called TIA, which enables a passive global adversary to accurately infer the traffic pattern in an anonymous MANET without compromising any node. As the first work of its kind, TIA works on existing on-demand anonymous MANET routing protocols. Detailed simulations show that TIA can infer the traffic pattern with an accuracy as high as 95%. Our results highlight the necessity for cross-layer designs to defend a MANET against traffic analysis.

The proposed system present a novel Traffic Inference Algorithm, called TIA, which allows a global adversary to accurately infer the MANET traffic pattern despite the use of existing anonymous on-demand routing protocols. TIA first exploits the overheard routing frames for flow recognition and then traces each flow in rounds based on the data-frame inter-arrival times. Given $\lambda > 2$ consecutive frames of a flow, the vector of $\lambda - 1$ corresponding frame inter-arrival times on any link is highly correlated with that of the next link along the flow path. This enables the adversary to iteratively derive the hidden flows and thus the traffic pattern from overheard MAC frames without prior knowledge of the inter-arrival time distribution of any flow.

In addition, GSTPDS does not need the signal detectors to be able to precisely locate the signal source. They are only required to determine which super node (region) the signals are sent from. Moreover, in STPDS, the actual receiver of a point-to-point transmission is not identifiable among all the potential receivers within the sender's transmitting range. This inaccuracy can be mitigated in GSTPDS because most potential receivers of a packet will be contained within one or a few super nodes.

IV. Traffic Pattern Discovery

In this module, Source / Destination Probability Distribution' and End-to-End Link Probability Distribution' are found out.

A. Source/Destination Probability Distribution

In this module, the actual source and destination probability Distribution are denoted respectively, as two vectors S =(s(1), s(2), ... , s(N) and D =(d(1), d(2), ... , d(N)), where s(i) and d(i) (i = 1 to N) represent the probability for node i to be an actual source and destination, respectively. Note that if the total number of source nodes is m, then $\sum_{i=1}^{N} s(i) = m$ for S. However, only care about the relative order among all possibilities (to know which nodes are more possible to be the actual sources) but not the total number m, always assume m = 1.

During the distribution finding, the vector space similarity (or cosine similarity) of two vectors V and U is defined as follows:

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

By introducing the vector space similarity assessment, ensure that, two nodes with higher probability to be neighbors (relays of each other) have less impact on each

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

other's source/destination probability distribution.

B. End-to-End Link Probability Distribution

This module derives a probability distribution matrix P =(p(i, j))NxN, in which each entry p(i, j) represents the probability of the i □ j linkability (i.e., node i and node j are a pair of actual source and destination). Again, note that only the relative order among these entries is of interest, since aim at discovering the most possible communication links. As described above, the probability for node i to be a destination depend on two factors: the traffic from each node j to node i and node j's probability to be a source. Suppose j - i is an actual source-destination pair. If set the total traffic coming out from j to zero, the probability for i to be a destination will decrease. Similarly, set the incoming traffic to node i to zero, the probability for node j to be a source will also decrease.

Identify a source-destination (S-D) pair by evaluating the significance of the probability reduction due to the elimination of the traffic sent by the source or received by the destination. For instance, in the example scenario shown in Fig. 1, to identify the most possible destination of node 1, erase all traffic sent by node 1 from the point-to-point traffic matrices, base on which compute the destination probability distribution D^-. By comparing D^- with D (obtained using the original point-to-point matrices), find out the node whose destination probability drops most significantly due to elimination of the traffic sent by node 1. This node is most possible to be the destination of node 1. Functions are used to remove the traffic sent by node i. as well as to remove the traffic received by node j.

V. EXPERIMENTAL RESULTS

The following Table describes experimental result for existing system successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are shown

The following Figure describes experimental result for existing system successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are shown

TABLE 2
NODE TRANSMISSION ANALYSIS

SERIAL NO	NUMBER OF TIME SLOT (M)	RATIO OF SUCCESSIVE TRANSMISSION NODE
1	10	0.48
2	20	0.57
3	40	0.66
4	60	0.72
5	80	0.77
6	100	0.83
7	120	0.89
8	140	0.92

TABLE I
NODE TRANSMISSION ANALYSIS

SERIAL NO	NUMBER OF TIME SLOT (M)	RATIO OF SUCCESSIVE TRANSMISSION NODE
1	10	0.43
2	20	0.52
3	40	0.61
4	60	0.69
5	80	0.74
6	100	0.80
7	120	0.86
8	140	0.90
9	150	0.93
10	160	0.97

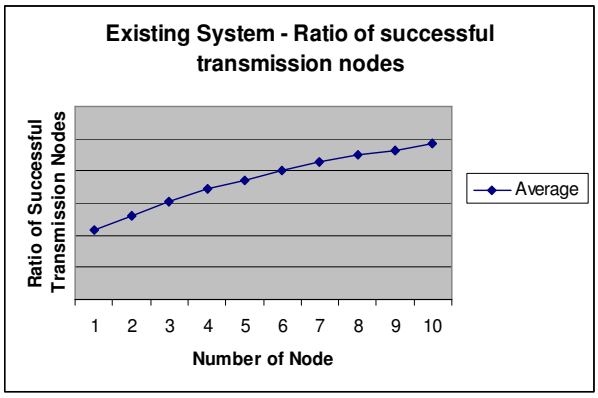

Fig.3.Ratio of successive transmission node analysis

Experimental result for proposed system successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are

The following Figure describes experimental result for proposed system successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are shown

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com
International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

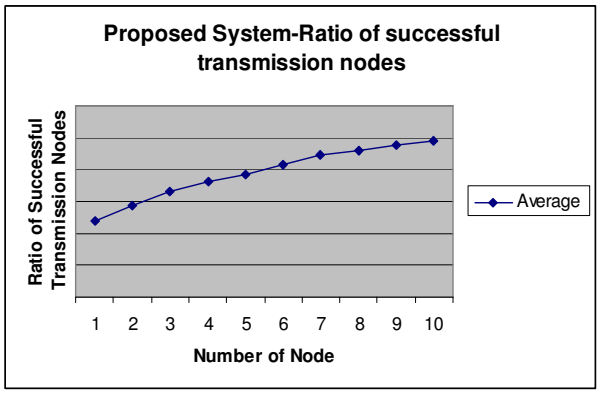

Fig.4 .Ratio of successive transmission node analysis

Periodic query scheduling for data aggregation with minimum delay under various wireless interference models. Time scheduling on a single frequency channel with the aim of minimizing the number of time slots required (schedule length) to complete converge-cast is considered. Next, scheduling with transmission power control is combined to mitigate the effects of interference, and show that while power control helps in reducing the schedule length under a single frequency, scheduling transmissions using multiple frequencies is more efficient. Lower bounds on the schedule length are given when interference is completely eliminated, and propose algorithms that achieve these bounds.

The following Table describes experimental result for differences existing system (EM) and proposed system (PM) successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are shown

TABLE 3
NODE TRANSMISSION ANALYSIS

S.NO	NUMBER OF TIME SLOT (M)	RATIO OF SUCCESSIVE TRANSMISSION NODE	
		EXISTING	PROPOSED
1	10	0.43	0.48
2	20	0.52	0.57
3	40	0.61	0.66
4	60	0.69	0.72
5	80	0.74	0.77
6	100	0.80	0.83
7	120	0.86	0.89
8	140	0.90	0.92
9	150	0.93	0.95
10	160	0.97	0.98

The following diagram provides the messages to all participants, each bit is sent around a second time by the participant at the end of the loop. Such an adapted ring requires, on average, a fourfold increase in bandwidth over the obvious traceable protocols in which messages travel only halfway around on average before being taken off the ring by their recipients. Rings differ from the dinner table in that several bit-transmission delays may be required before all the outputs of a particular round are known to all participants; collisions are detected only after such delays. Efficient use of many other practical communication techniques requires participants to group output bits into blocks.

Figure describes experimental result for differences existing system (EM) and proposed system (PM) successive transmission node analysis. The table contains number of time slot interval and given time interval to calculate average numbers of send transmission node details are shown

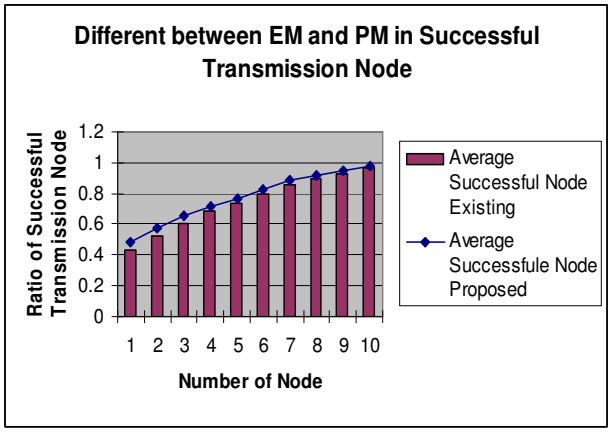

Fig. 5. Comparison of EM and PM in Successful Transmission Node

VI. CONCLUSION

STARS are basically an attacking system, which only needs to capture the raw traffic from the PHY/MAC layer without looking into the contents of the intercepted packets. From the captured packets, STARS constructs a sequence of point-to-point traffic matrices to derive the end-to-end traffic matrix, and then uses a heuristic data processing model to reveal the hidden traffic patterns from the end-to-end matrix. The empirical study demonstrates that the existing MANET systems can achieve very restricted communication anonymity under the attack of STARS. In addition, the STPDS is extended as GSTPDS which divides the entire network into multiple regions geographically; and deploys sensors along the boundaries of each region to monitor the cross-component traffic. Also it treats each region as a super node and use STPDS to figure out the sources, destinations, and end-to-end communication relations.

ISSN (ONLINE): 2454-9762
ISSN (PRINT): 2454-9762
Available online at www.ijarmate.com

IJARMATE
Your ubi-MATE Research Paper !!!

International Journal of Advanced Research in Management, Architecture, Technology and Engineering (IJARMATE) Vol. 2, Issue 3, March 2016

VII. REFERENCES

[1] Kong J., Hong X. and Gerla M., *"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.

[2] Benny Chor, Oded Goldreich, Eyal Kushilevitz and Madhu Sudan, *"Private information retrieval"*, Journal of the ACM 45(6), 965–981, 1998.

[3] Choi H., McDaniel P. and La Porta T.F., *"Privacy preserving communication in MANETs"*, in IEEE SECON'07 San Diego CA, pp. 233–242, Jun.2007.

[4] Hu Y.C., Perrig A. and Johnson D.B., *"Ariadne: A secure on-demand routing protocol for ad hoc networks"*, ACM MobiCom Atlanta GA, Sep. 2002.

[5] Kong j. and Hong X., *"ANODR: Anonymous on demand routing with untraceable routes for mobile ad-hoc networks"*, ACM MobiHoc'03 Annapolis MD, pp. 291 – 302, Jun. 2003.

[6] Liu Y., Zhang R., Shi J. and Zhang Y., *"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, 2010.

[7] Oded Goldreich and Rafail Ostrovsky, *"Software protection and simulation on oblivious RAMs"*, Journal of the ACM 43(3), pp. 431–473, 1996.

[8] Perkins C., Belding-Royer E. and Das S., *"Ad hoc on-demand distance vector (AODV) routing"*, RFC 3561, July 2003.

[9] Qin Y. and Huang D., *"OLAR: On-Demand Lightweight Anonymous Routing in MANETs"*, Proc. Fourth Int'l Conf. Mobile Computing and Ubiquitous Networking (ICMU '08), pp. 72-79, 2008.

[10] Rabin M., *"How to exchange secrets by oblivious transfer Technical Report Technical Memo TR-81"*, Aiken Computation Laboratory Harvard University, 1981.

[11] Ronald Cramer, *"Introduction to secure computation"*, Lectures on data security: modern cryptology in theory and practice Lecture Notes in Computer Science, vol. 1561, pp. 16–62. Springer 1999.

[12] Sanzgiri K., Dahill B., Levine B., Shields C. and Royer E., *"A secure routing protocol for ad hoc networks"*, IEEE ICNP'02 Paris France, Nov. 2002.

[13] Seys S. and Preneel P., *"ARM: Anonymous Routing Protocol for Mobile Ad Hoc Networks"*, Proc. IEEE 20th Int'l Conf. Advanced Information Networking and Applications Workshops (AINA Workshops '06), pp. 133-137, 2006.

[14] eys S. and Preneel B., *"Arm: Anonymous routing protocol for mobile ad hoc networks"*, IEEE AINA, 2006.

[15] Troncoso C., B. Gierlichs, B. Preneel and I. Verbauwhede, *"Perfect Matching Disclosure Attacks"*, Proc. Eighth Int'l Symp. Privacy Enhancing Technologies, pp. 2-23, 2008.

[16] Waidner M., *"Unconditional sender and recipient untraceability in spite of active attacks"*, Advances in Cryptology – Eurocrypt '89 of Lecture Notes in Computer Science, vol. 434, Springer-Verlag 1989.

[17] Wang W., S. Chen and S. Jajodia, *"Network Flow Watermarking Attack on Low-Latency Anonymous Communication Systems"*, Proc. IEEE Symp. Security and Privacy, pp. 116-130, 2007.

[18] Zhang Y., Liu W., Lou W. and Fang Y., *"MASK: Anonymous On-Demand Routing in Mobile Ad Hoc Networks"*, IEEE Trans. Wireless Comm., vol. 5, pp. 2376-2385, Sept. 2006.