A Novel Attack Tree Based Risk Assessment Approach for Location Privacy Preservation in the VANETs

Dandan Ren and Suguo Du
Antai College of Economics & Management
Shanghai Jiao Tong University
Shanghai, China
{rendandan, sgdu}@sjtu.edu.cn

Haojin Zhu
Department of Computer Science & Engineering
Shanghai Jiao Tong University
Shanghai, China
zhu-hj@cs.sjtu.edu.cn

Abstract—**Even though emerging as a promising approach to increase road safety, efficiency and convenience, Vehicular Ad hoc Networks (VANETs) pose many new research challenges, especially on the aspect of location privacy. The existing literatures focus on preventive techniques to achieve location privacy protection, however the location privacy risk assessment receives less attention. In this paper, we introduce a novel risk assessment method to evaluate the security risk of VANET's privacy based on attack tree. The proposed scheme provides a general analysis framework to estimate the degree that a certain threat might bring to the VANETs. We also use the constructed attack tree to identify possible attack scenarios that an attacker may launch towards the privacy preserving system in VANETs, which is expected to further improve the system security.**

Keywords-Location privacy; Attack tree; VANET; Risk assessment;

I. INTRODUCTION

Vehicular ad hoc networks (or VANETs) are self-organized networks designed for communication among vehicles [1]. In VANET, each vehicle is equipped with an On Board Unit, by which vehicles are able to communicate with each other as well as Road Side Unit. VANETs are expected to support a wide range of promising applications such as location based services. However, the broadcast nature of the wireless medium allows an adversary to eavesdrop on the communications containing node identifiers, and to estimate the locations of the communicating nodes with an accuracy that is sufficient for tracking the nodes. For example, in [2], it is reported that the adversary can even approximately derive the drivers' family address and their workplaces with the collection of location traces every day. Due to these reasons, location privacy threat has been well recognized as one of the major security threats for VANETs and has attracted a lot of interest recently [3].

The existing research on VANET privacy mainly focuses on the preventive techniques including: pseudonym based approaches [4], group signature based techniques [5] or mix-zone based approach [3][6][7]. The basic idea of the preventive techniques is introducing the privacy protection techniques to prevent the compromise of user location privacy. However, the preventive schemes may face the challenges that it can only address the expected security vulnerabilities while fails to deal with the unexpected privacy threats. Furthermore, from a system point of view, a comprehensive yet well-defined security evaluation enables the system administrator to identify the most critical security threats and attack strategies, which are more than important for the overall success of VANET deployment. Some reported studies investigate the possible security vulnerabilities. In [8], it proposed a general method to assess risk, which is not especially for location privacy in VANETs. In [9], it identifies attacks threatening the privacy of users based on simple cost estimation. However, they fail to give a quantity risk analysis from a system point of view for location privacy.

In this study, we propose a novel risk assessment approach for location privacy preserving in VANETs based on the attack tree based approach. Attack tree based risk analysis leverages tree based method to model and analysis the risk of the system and identify the possible attacking strategies the adversaries may launch. With the help of the attack tree model, it is convenient to analyze the capability of the attack source and estimate the degree or the impact a certain threat that might bring to the system. Due to these features, in this paper, we take advantage of attack tree based approach to identify all the possible threats. And we further calculate the total probability of reaching attack goal on the basis of the attack tree. According to the quantitative result, the decision maker of the system can decide which protection measure should be adopted.

The reminders of this paper are organized as follows: in section II, we introduced the attack tree method and present how to build the attack tree. In section Ⅲ, we assigned values to leaf nodes and calculated the system's risk. In section Ⅳ, to estimate the most likely method an attacker may choose, we carried out an analysis of attack scenarios on the basis of attack tree. At last, in section Ⅴ, the conclusion was made.

This work was supported in part by National Natural Science Foundation of China (Grant No. 61003218 and 70971086), Doctoral Fund of Ministry of Education of China (Grant No.20100073120065).

978-1-61284-233-2/11/$26.00 ©2011 IEEE

II. Attack Tree Model for VANET Privacy

A. System Model

Communication in VANET is divided into two parts: vehicle to infrastructure and vehicle to vehicle. The followings are some assumptions about the network [7]:

- Each vehicle has its own communication equipment OBU (On Board Unit), which enables the vehicles to communicate with others as well as the Roadside Units (RSUs).
- It is assumed that there is a trusted third party called Certificate Authority, like transportation authority within the network to take charge of the network's security and privacy issues. Each vehicle becomes a legitimate node of the network until it registers at CA
- The CA disseminates each node with a single identity as well as a set of pseudonyms after it verified the validity of the node's identity.
- A node changes its pseudonym at certain intervals for the privacy preservation. Expired pseudonym is directly removed from the vehicle's storage media and CA is responsible for the issuance of new pseudonyms if a node uses up all of its pre-download pseudonyms.
- Each node automatically broadcasts its location, velocity and other special information to its neighbors at fixed intervals.
- Vehicles have enough power to install and run personal firewall or other antivirus software to protect it from malicious programs like worms and viruses spread among wireless network.

B. Building attack tree

An attack tree can be simply described as an analytical technique which has tree structure [10]. In an attack tree, its top event, which is also called attack goal, is an attacker's final desired purpose. That which event will be taken as attack goal depends on your aim of analysis. After the attack goal is identified, the system is analyzed in the context of its environment and operation to find all credible events that directly result in the attack goal's occurrence [11]. According to the logical relationship among them, those events are linked with an OR-gate or AND-gate. The OR-gate is used to show that the output event occurs only if one or more of the input events occur. The AND-gate shows that the output event occurs only if all the input events occur. If one of those events cannot be divided further, it is a leaf node. Otherwise, those events are gate nodes that are treated as sub-goals separately and be divided continually until all the events become leaf nodes. Whether an event can be proceeded down is determined by the depth of analysis as well as one's knowledge about the system. That is to say, an event may be a leaf node in an attack tree, but a gate node in another one.

In VANET system, we set *leakage of location privacy* as the attack goal, which is denoted by G. In the follows, we perform a top-to-down strategy to build the attack tree, which starts from G. The overall construction process is performed with a *divide-and-conquer* manner. The intermediate causes of the G are: direct communication, eavesdropping, stealing and illegal disclosure, which are respectively marked by M_1, M_2, M_3, and M_4. The mission objective can be achieved if any of the four components is reached. Now we identify the four intermediate causes as sub-goals, and it is necessary to determine their immediate cause or causes separately. We take the left two sub-trees in Figure 1 for illustration, while the other sub-trees are constructed with similar rules.

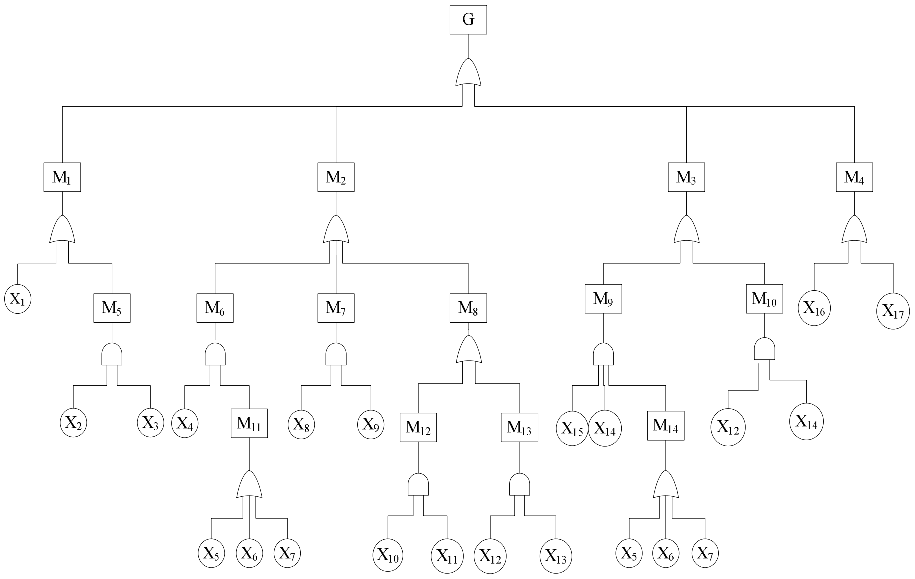
In the most left tree, there are two possibilities to achieve sub-goal "direct communication (M_1)".

- "Inquiry (X_1)": an attacker communicates with a target node with real identity of itself, and then inquiries the node's location privacy. This way applies only when a node has low sensitivity on its privacy.
- "Cheating (M_5)": an attacker impersonates as someone else that the target node trusts, and gets privacy of the target node by communicating with it. Since nodes which have close relationship with a target node such as friends, colleagues or service providers are supposed to be reliable for most nodes, this kind of attack succeeds with higher opportunity.

Therefore, the sub-goal "direct communication (M_1)" can arise from two events, "inquiry (X_1) "or "cheating (M_5)". Now we are ready to seek out the immediate causes for the new sub-goal "cheating (M_5)", which appears as intersection of two events: "finding vulnerabilities in the system's authentication mechanism X_2 " and "making fake identity X_3".

We now continue the analysis by focusing our attention on event "eavesdropping (M_2)". To get a node's privacy by "eavesdropping (M_2)", it is necessary to carry out "physical layer eavesdropping (M_6)", or "MAC layer eavesdropping (M_7)", or "application layer eavesdropping (M_8)". We identified M_6, M_7, and M_8 as intermediates which will be analyzed as below.

- Physical layer eavesdropping (M_6): This mission can be achieved if two tasks were accomplished in a row: "dismantle wiretap-proof device X_4" and "install wiretap device M_{11}". There are three ways for an attacker to successfully "install wiretap device M_{11}": "pretend to be a service provider X_5", or "disrupt the car's anti-theft system X_6", or "make use of the owner's carelessness X_7".
- MAC layer eavesdropping (M_7): The attack "MAC layer eavesdropping (M_7)" can be launched through two joint atomic attacks: "protocol vulnerability analysis X_8" and "reset its own configuration X_9".
- Application layer eavesdropping (M_8): In application layer, an attacker can choose from two aspects to



G	leakage of location privacy	X_2	finding vulnerabilities in the authentication mechanism
M_1	direct communication	X_3	make fake identity
M_2	eavesdropping	X_4	dismantle wiretap-proof device
M_3	stealing	X_5	being a service provider for cars
M_4	illegal disclosure	X_6	disrupt a car's anti-theft system
M_5	cheating	X_7	make use of the owner's carelessness
M_6	physical layer eavesdropping	X_8	protocol vulnerability analysis
M_7	MAC layer eavesdropping	X_9	reset its own configuration
M_8	Application layer eavesdropping	X_{10}	obtain signal receiver
M_9	physical theft	X_{11}	analyze the adopted pseudonym mechanism's weakness
M_{10}	malicious code theft	X_{12}	climbing over the network's firewall
M_{11}	install wiretap tool	X_{13}	be familiar with wireless network's weak security feature
M_{12}	eavesdropping pseudonyms	X_{14}	decipher the encrypted file
M_{13}	running eavesdropping software	X_{15}	disrupt the function of removing data from remote control
M_{14}	stealing the car	X_{16}	purchase privacy from third party
X_1	inquiry	X_{17}	leakage from official department

FIGURE 1 ATTACK TREE MODEL FOR VANET PRIVACY

eavesdrop: “eavesdropping pseudonyms M_{12}” or “running eavesdropping software M_{13}”. Each of the two is further decomposed into sub-components. For “eavesdropping pseudonyms M_{12}”, it is essential for an attacker to “obtain signal receiver X10” and “analyze the adopted pseudonym mechanism's weaknesses X11”. While for “running eavesdropping software M13”, after “climbing over the network's firewall X12” an attacker needs to “be familiar with wireless network's weak security feature X13” and then compromise the target node by planting eavesdropping program.

Sub-components of “stealing M_3” and “illegal disclosure M_4” are decomposed in a similar way. Note that, the attack tree described above is only an example to capture the possible ways launched by the adversaries. As a general framework, in practice, an attack tree could accommodate more attack sub-trees by considering more attack strategies of the adversaries. We believe that the recent research progress on VANET security and privacy will also benefit the construction of trees [14-22].

III. RISK ASSESSMENT

For the limitation of resources, an attacker has to take into account of many aspects including the possibility to succeed, attack cost, technique difficulty, risk of being detected and so on. In this paper, we calculate the total probability of reaching the attack goal by assigning leaf nodes three attributes: attack cost, technical difficulty and the probability to be discovered, which are denoted as c_L, d_L and s_L respectively. The grade level standards are given in Table Ⅰ.

TABLE I. GRADE STANDARD

Attack cost/ten thousands		**Technical difficulty**		**Probability to be discovered**	
Attack cost c_L	***grade***	***Technical difficulty*** d_L	***grade***	***Probability to be discovered*** s_L	***grade***
>10	5	quite difficult	5	quite difficult	1
6-10	4	difficult	4	difficult	2
3-6	3	mediate	3	mediate	3
0.5-3	2	simple	2	simple	4
<0.5	1	quite simple	1	quite simple	5

The specific assignment of each node's attribute requires knowledge of implementation details of the system including protocols, hardware, operating system as well as attack software and tools. Since the main concern of this paper is on the new evaluation method in location privacy, we try to estimate values of c_L, d_L and s_L for each leaf node according to following rules: (values that are assigned to each leaf node are listed in Table Ⅱ)

- An attacker could launch an attack on any layer of the system. The higher the attacked layer is, the more difficult for the attacker, and the lower the probability of success. In this condition, we can sort those layers according to the difficulty of compromise in the following order: application layer > transmission layer > routing layer > MAC layer > physical layer.
- Cost of physical vehicle and labor assistant is higher compared with that of analyzing protocol or mechanism's vulnerabilities.
- Powerful nodes, for instance traffic offices or service providers, are capable of taking more protecting behaviors, so these nodes are more difficult to be compromised than common nodes.

TABLE II. ATTRIBUTE VALUES FOR EACH LEAF NODE

Leaf node	**Attribute**			**Occurrence probability**
	Attack cost	***Technical difficulty***	***Probability to be discovered***	
X_1	3	2	5	0.069
X_2	3	4	2	0.072
X_3	2	1	4	0.117
X_4	1	2	2	0.133
X_5	5	1	5	0.093
X_6	2	3	4	0.072
X_7	1	2	3	0.122
X_8	3	3	1	0.111
X_9	4	2	1	0.117
X_{10}	2	4	2	0.083
X_{11}	3	4	3	0.061
X_{12}	2	3	2	0.089
X_{13}	2	5	1	0.113
X_{14}	1	2	4	0.117
X_{15}	4	5	5	0.043
X_{16}	4	4	3	0.056
X_{17}	5	2	5	0.060

After value assignment for leaf nodes, multi-attribute utility theory is adopted to transfer these three attributes into attackers' utility value [12], which is the occurrence probability of a leaf node. The following is a formula we applied to calculate the utility of each leaf node:

$$P_L = w_1 \times u_1(c_L) + w_2 \times u_2(d_L) + w_3 \times u_3(s_L)$$

Where $u_1(c_L)$, $u_2(d_L)$, $u_3(s_L)$ represent the utility functions of c_L, d_L and s_L separately , and their values fall into the interval of [0,1]; w_1, w_2, w_3 are the utilities' correspondent weights, where $w_1+w_2+w_3=1$.

For the convenience of calculation, in this paper we define that $w_1=w_2=w_3=1/3$. In order to get the occurrence probability of each leaf node, we also need to determine the utility function. Since all the three attributes are inversely proportional to their respective utility value, we suppose that the three utility functions $u_1(c_L)= u_2(d_L)=u_3(s_L)=u(x)=c/x$ (where, the value of parameter c is chosen as 0.2 in this paper to guarantee that $P_L \in [0, 1]$ even if w_1, w_2, w_3 differs). Then the occurrence probability of each leaf node can be calculated (see the rightmost column in Table Ⅱ) by using the utility functions combined with the attribute values assigned to leaf nodes.

So as to calculate the total probability of reaching the attack goal, the attack tree is transferred to a BDD [13] (Binary Decision Diagram). According to probability counting rules in a BDD, we get that the total probability of reaching the attack goal is 0.239.

IV. ATTACK SCENARIOS

An attack scenario is a set of leaf nodes, in which only the occurrence of all the leaf nodes could reach the attack goal, that is to say the goal will not be realized if one of the leaf nodes does not occur. An attack scenario is the real attack path that an attacker considers. Once the attack scenarios have been known, we could calculate their probabilities of occurrence, and then compare them to find out the attack scenario that the malicious may launch most likely. Suppose an attack scenario is denoted as:

$$S_i = (X_{i_1}, X_{i_2}, \cdots, X_{i_n})$$

Then the probability of an attack scenario is:

$$P(S_i) = P(X_{i_1}) \times P(X_{i_2}) \times \cdots \times P(X_{i_n}) \tag{1}$$

We now adopt Boolean algebra method to get all the attack scenarios for our attack tree.

$$\begin{aligned}
&M_1=X_1+M_5=X_1+(X_2*X_3)\\
&M_2=M_6+M_7+M_8\\
&=(X_4*M_{11})+(X_8*X_9)+(M_{12}+M_{13})\\
&=(X_4*(X_5+X_6+X_7))+(X_8*X_9)+(X_{10}*X_{11}+X_{12}*X_{13})\\
&M_3=M_9+M_{10}\\
&=(X_{15}*X_{14}*M_{14})+(X_{12}+X_{14})\\
&=(X_{15}*X_{14}*(X_5+X_6+X_7))+(X_{12}*X_{14})\\
&M_4=X_{16}+X_{17}\\
&G=M_1+M_2+M_3+M_4\\
&=X_1+(X_2*X_3)+(X_4*(X_5+X_6+X_7))+(X_8*X_9)+(X_{10}*X_{11}+X_{12}*X_{13})\\
&+(X_{15}*X_{14}*(X_5+X_6+X_7))+(X_{12}*X_{14})+X_{16}+X_{17}\\
&=X_1+X_2*X_3+X_4*X_5+X_4*X_6+X_4*X_7+X_8*X_9+X_{10}*X_{11}\\
&+X_{12}*X_{13}+X_{15}*X_{14}*X_5+X_{15}*X_{14}*X_6+X_{15}*X_{14}*X_7+X_{12}*X_{14}+X_{16}+X_{17}
\end{aligned}$$

From the formula of G, we can see that there are fourteen attack scenarios to achieve the attack goal, and they respectively are $\{X_1\}$, $\{X_2, X_3\}$, $\{X_4, X_5\}$, $\{X_4, X_6\}$, $\{X_4, X_7\}$,

{X_8, X_9}, {X_{10}, X_{11}}, {X_{12}, X_{13}}, {X_5, X_{14}, X_{15}},{X_6, X_{14}, X_{15}},{X_7, X_{14}, X_{15}},{X_{12}, X_{14}},{X_{16}},{X_{17}}. In the first scenario {X_1}, it means an attacker can get the target's privacy by just launch the atomic attack X_1, while in the second scenario {X_2, X_3}, an attacker requires to compromise the system's authentication mechanism X_2 as well as making fake identity X_3. Therefore, we could use the equation (1) to calculate probability of occurrence for each attack scenarios (In Table Ⅲ).

TABLE III. PROBABILITY OF ATTACK SCENARIOS

Attack Scenarios	Leaf nodes in each scenarios	Probabilities
S_1	X_1	0.0690
S_2	X_2, X_3	0.0084
S_3	X_4, X_5	0.0123
S_4	X_4, X_6	0.0096
S_5	X_4, X_7	0.0162
S_6	X_8, X_9	0.0130
S_7	X_{10}, X_{11}	0.0051
S_8	X_{12}, X_{13}	0.0101
S_9	X_5, X_{14}, X_{15}	0.0005
S_{10}	X_6, X_{14}, X_{15}	0.0004
S_{11}	X_7, X_{14}, X_{15}	0.0006
S_{12}	X_{12}, X_{14}	0.0104
S_{13}	X_{16}	0.0560
S_{14}	X_{17}	0.0600

From the Table Ⅲ, we find that the attack scenario S_1 is the most likely happened attack path, so to protect the system from attack, it is necessary to first keep close eyes on it and take correspondent location protection measures.

V. CONCLUSIONS

Location privacy in VANET is getting more concerns of the world. However, the study on the risk assessment of location privacy hasn't been paid much attention. This paper in the first time adopted attack tree methodology to quantitatively assess the risk of location privacy in VANET from the system's perspective. In this paper, we build an attack tree with the leakage of location privacy information as attack goal. Also the total possibility of reaching attack goal is calculated on the basis of the attack tree. At last, according to the attack scenario analysis, we find out the most likely path that an attacker may use. The future work could be to build defense tree of the system's location privacy, by which the system's overall defense capability could be evaluated. These works also could provide theoretical support for decision-makers to choose adequate protection for the VANET system.

REFERENCES

[1] C. Cuyu, X. Yong, S. Meilin, "Development and status of vehicular ad hoc networks", *Journal on Communications*, no. 28, vol. 17, pp. 116-126, 2007.

[2] J. Krumm, "Inference attacks on location tracks", in Proc. of *the 5th International Conference on Pervasion Computing*, Toronto, Ontario, Canada. Springer-Verlag, 2007, 127-143.

[3] C. Troncoso, G. Danezis, "The Bayesian traffic analysis of mix networks", in Proc. of *the 16th ACM Conference on Computer and Community Security. ACM CCS'09*,369-379, 2009.

[4] T. Leinmuller, E. Schoch and C. Maihofer, "Security requirements and solution concepts in vehicular ad hoc networks", in Proc. of *the 4th Annual Conference on Wireless Demand Network Systems and Services*, 2007.

[5] X. Lin, X. Sun, P.-H. Ho and X. Shen, "GSIS: A Secure and Privacy-Preserving Protocol for Vehicular Communications", *IEEE Transactions on Vehicular Technology*, vol. 56, no. 6, pp. 3442-3456, 2007.

[6] J. Freudiger, M. Raya, M. Felegyhazi, "Mix-zones for location privacy in vehicular networks", in Proc. of *ACM WiN-ITS*'07, 2007.

[7] J. Freudiger, R. Shokri, and J.-P. Hubaux, "On the optimal placement of mix zones", in Proc. of *the 9th International Symposium on Privacy Enhancing Technologies*, Springer-Verlag, 2009, 216-234.

[8] M. Gerlach, "Assessing and improving privacy in VANET", in Proc. of *Fourth Workshop on Embedded Security in Cars (ESCAR)*, November 2006.

[9] Risk-Based Systems Security Engineering: Stopping Attacks with intention, *IEEE Security & Privacy*, 2004, 59-62

[10] Schceier B. Attack trees. *Dr. Dobb's Journal*, 1999, 24(12):21-29

[11] William E. Vesely, N. H. Roberts. *Fault tree handbook*,1981

[12] R. A. Kemmerer, "Covert flow trees: a visual approach to analyzing covert storage channels", *IEEE Transaction on Software Engineering*, no.17, vol.11, pp.1166-1185, 1991.

[13] Y. Sun. A research on variable ordering methods of binary decision diagram. Shanghai: Shanghai JiaoTong University, 2008.

[14] H. Zhu, R. Lu, X. Lin and X. Shen, "Security in Service-Oriented Vehicular Networks", *IEEE Wireless Communication Magazine*, vol. 16, no.4, pp. 16-22, August, 2009.

[15] H. Zhu, X. Lin, R. Lu, Y. Fan, and X. Shen, "SMART: A Secure Multi-Layer Credit based Incentive Scheme for Delay-Tolerant Networks", *IEEE Trans. on Vehicular Technology*, vol.58, no. 8, pp. 4628-4639, 2009.

[16] H. Zhu, X. Lin, R. Lu, P.H. Ho, and X. Shen, "SLAB: Secure Localized Authentication and Billing Scheme for Wireless Mesh Networks", *IEEE Trans. on Wireless Communications*, vol. 17, no. 10, Oct. 2008.

[17] Q. Han, S. Du, D. Ren and H. Zhu, SAS: A Secure Data Aggregation Scheme in Vehicular Sensing Networks, in Proc. Of *International Conference on Communications (IEEE ICC'10)*, Cape Town, South Africa, May 23-27, 2010.

[18] R. Lu, X. Lin, and X. Shen, "SPRING: A Social-based Privacy-preserving Packet Forwarding Protocol for Vehicular Delay Tolerant Networks", Proc. *IEEE INFOCOM'10*, San Diego, California, USA, March 14 - 19, 2010.

[19] R. Lu, X. Lin, H. Zhu, P.H. Ho and X. Shen, "ECPP: Efficient Conditional Privacy Preservation Protocol for Secure Vehicular Communications", Proc. IEEE *INFOCOM'08*, Phoenix, AZ, USA, April 14-18, 2008.

[20] R. Lu, X. Lin, H. Zhu, and X. Shen, "SPARK: A New VANET-based Smart Parking Scheme for Large Parking Lots", Proc. IEEE *INFOCOM'09*, Rio de Janeiro, Brazil, April 19-25, 2009.

[21] H. Zhu, X. Lin, R. Lu, X. Shen, D. Xing and Z. Cao. An Opportunistic Batch Bundle Authentication Scheme for Energy Constrained DTNs. *The 29th IEEE International Conference on Computer Communications (INFOCOM 2010)*, San Diego, California, USA, March 14-19, 2010.

[22] C. Zhang, R. Lu, X. Lin, P.-H. Ho and X. Shen. An Efficient Identity-based Batch Verification Scheme for Vehicular Sensor Networks. *The 27th IEEE International Conference on Computer Communications (INFOCOM 2008)*, Phoenix, Arizona, USA, April 15-17, 2008.