Experimental Privacy Analysis and Characterization for Disconnected VANETs

Chibueze P. Anyigor Ogah(✉), Haitham Cruickshank, Philip M. Asuquo, Ao Lei, and Zhili Sun

5G Innovation Centre (5GIC), Institute for Communication Systems, University of Surrey, PATS Driveway, Guildford, Surrey GU2 7XS, UK
{c.anyigorogah,h.cruickshank,p.asuquo,a.lei,z.sun}@surrey.ac.uk

Abstract. Intelligent Transport Systems (ITS) are special applications of Vehicular Ad-hoc Networks (VANETs) for road safety and efficient traffic management. A major challenge for ITS and VANETs in all its flavours is ensuring the privacy of vehicle drivers and the transmitted location information. One attribute of ITS during its early roll-out stage especially in rural areas and challenged environments is low vehicle density and lack of end-to-end connectivity akin to the attribute of Vehicular Delay Tolerant Networks (VDTNs). This means that contact duration between network entities such as vehicles and road-side units (RSUs) are short-lived. Three popular solutions are the use of pseudonyms, mix-zones, and group communication. Privacy schemes based on the mix-zone technique abound for more conventional VANETs. A critical privacy analysis of such scenarios will be key to the design of privacy techniques for intermittent networks. We are not aware of any work that analyse the privacy problem in intermittent VANTEs. In this paper, we add our voice to efforts to characterize the privacy problem in disconnected VANETs.

Keywords: Anonymity · Evaluation · ITS · Privacy · VANETs · Vehicular delay tolerant networks · VDTN

1 Introduction

With the application of VANETs for Intelligent Transport Systems (ITS), enhancing road safety and traffic management becomes more effective and cost efficient. Already, there are a handful of driver-less car projects all over the world, with examples such as the Google Car project [1]. Self-driving cars have been tested in Europe when a fleet of trucks made a voyage journey across the continent from Rotterdam with no incidents [2]. Google Cars are also driving across California in pilot test-drives on a regular basis with only one incident of error on the part of the vehicles reported so far. The IEEE defines ITS as those systems utilizing synergistic technologies and systems engineering concepts to develop and improve transportation systems of all kinds. These include applications that depend on vehicle-to-vehicle (V-to-V) and vehicle-to-infrastructure

I. Otung et al. (Eds.): WiSATS 2016, LNICST 186, pp. 119–129, 2017.
DOI: 10.1007/978-3-319-53850-1_13

(V-to-R) communication for road safety and improved traffic management [3]. A variety of ITS exist and have been well researched on. Traditional disconnected VANETs such as Delay Tolerant Networks (DTNs) do not make use of infrastructure support such as Road-Side Units (RSUs). However, recent efforts towards improving reliability and security have inspired the introduction of infrastructure [4], thereby creating such as flavors as Vehicular Delay Tolerant Networks [5].

Despite its advantages, there is yet no consensus on how to exactly address the key issues of security and privacy [6]. Privacy issues results from the fact that a malicious user can intercept the location information contained in safety messages to track a driver's location. In reality, tracking a vehicle is as good as tracking its driver or owner. Again, when these technologies are fully developed, there will be the problem of inadequate infrastructure to support thousands of vehicle in both urban and rural areas especially at the early stages of deployment. Hence, it will not be possible to conduct a thorough roll-out especially in rural areas due in part to lack of adequate infrastructure. Therefore, performance evaluation from a pilot-phase study will be compulsory to understand its viability for large-scale deployment. While a big chunk of the pilot study may concern bandwidth support for thousands of vehicle, there is also the need to understand and address security and privacy related issues. The above issues forms the crux of privacy problems in VANETs. Our focus on this paper is to add our voice towards characterizing and evaluating privacy in a disconnected VANET. This is a progressive effort towards proffering adequate solutions.

A variety of schemes have been proposed in literature to address privacy in VANETs. In the United States and Europe for instance, the Dedicated Short Range Communications/Wireless Access in Vehicular Environments (DSRC/WAVE) specifies the formats of Basic Safety Messages for ITS. Within the IEEE 802.9 family of standards, the IEEE 1609.2 makes provisions for location privacy but does not specify its modalities. Three popular Privacy-Enhancing Techniques (PETs) in VANETs include pseudonymous communication, mix-zones, and group communication [7–9]. A number of other schemes exist, most of which are indeed variants of those mentioned here; these include Silent Periods [10], Virtual Mix-Zones [11], and pseudonym management techniques such as [12]. In this work, we criticize group communication on the basis that it is difficult to easily to find group collaborators especially in light traffic situations such as in rural areas. Again, pseudonym-based privacy schemes especially in disconnected VANETs have a direct relationship with vehicle density, traffic load and generation, and mobility pattern [11,12]. The peculiar operating environment of VANETs for ITS especially in its pilot and early deployment phases can best be described as disconnected, hence even traditional PETs such as those mentioned above will fail [13]. For the same reasons, conventional mix-zones techniques are not efficient because they are limited by the number of vehicles that can potentially collaborate for pseudonym change around pseudonym changing spots [12]. While pseudonym and mix-zone techniques remains one of the most popular and documented PETs, more effort is needed in ensuring their effectiveness in disconnected environments. In order not to re-invent the wheel,

we leverage these established techniques and rather focus on discussing their application in disconnected networks.

To this end, we set out to characterize and analyse the privacy behavior in a disconnected VANETs. Our research here extends our initial idea in [13] for group-based communication in VDTNs. We characterize the mobility dynamics in our network and relation to privacy. Our analysis is based on a formal model and compared with existing literature. The rest of this paper is organized as follows. We describe our scheme and it's attributes in Sect. 2. In Sect. 3, we provide a summary of the problem we set out to solve and outline key definition of terms. Our detailed simulation and system analysis is provided in Sect. 4, and finally, we conclude and present our future work in Sect. 5.

2 Model Description

In this section, we present a detailed description of our model, assumptions and the attributes of the adversary.

2.1 Network Model

We consider disconnected VANET deployment in a rural area as described in [5,14] comprised of mobile vehicles, stationery RSUs, and a central administrative authority known as the Trusted Key Manager (TKM). Our network can be modelled as a directed multi-graph, $\mathcal{G} = (\mathcal{V},\mathcal{E})$ where $\mathcal{V}$ and $\mathcal{E}$ denote some fleet of vehicles and contact edges respectively. The RSUs act as stationary relay nodes that facilitate packet routing in addition to assisting with the security and privacy administration of the TKM. We only deploy RSUs at strategic locations which we regard as density zones as in [15].

2.2 Adversary Model

We consider a global passive adversary in our model. The adversary can isolate sections of the network and monitor communication and beacon messages exchanged between vehicles within each density zone to resolve vehicle and driver identity. We assume that the RSUs are trustworthy and tamper-proof while the vehicles are not. Hence, the vehicles can deviate in behaviour and act as adversaries (e.g. by reporting wrong location information). Examples of specific location privacy related attacks the adversary can execute include tracking and packet analysis attacks. To be able to execute packet analysis attacks, the adversary can delay the message delivery for a considerable amount of time while analysing it to divulge information regarding source and destination vehicles.

3 Problem Description and Definition of Terms

Our main objective is to characterize and analyse the privacy problem in our model in relation to the disconnectedness and mobility pattern of the network. We define some terms and describe the privacy problems in the following section.

3.1 Privacy Analysis

Our analysis is based on Shanon's information theory. We take into account the fact that location privacy depends on vehicle density [16]. The relationship between vehicle density and location privacy is easily understood from the point of view of changing pseudonyms - higher vehicle density means its a higher probability for a vehicle to find potential pseudonym collaborators. Since our network is disconnected, the density and mobility pattern of vehicles in the network and around the density zones will be key to how much privacy is achieved. The following usual definitions relate to the privacy analysis of our model.

3.1.1 Anonymity, Entropy, Anonymity Set, and Anonymity Duration

Anonymity. The anonymity of a vehicle, $V_i \in \mathcal{V}$ can be defined as a state of being unidentifiable among *k-1* other vehicles. Anonymity is usually related to the *unlinkability* property. Unlinkability is a term used to describe the notion that the adversary cannot link the vehicles identity, V_i to two actions say, $Actions_1$ and $Actions_2$ executed at different times t_1 and t_2 in relation the locations l_1 and l_2 where the actions took place with ease. A typical example of an action can be a vehicle changing pseudonyms or sending a message. This means that a vehicle cannot be linked to its identity for a duration of time due to its activity on the network. The IEEE 1609.2 measures location privacy using anonymity [17,18].

Anonymity Set. The anonymity set, AS is the average number of vehicles that are indistinguishable from *k-1* other vehicles from the privacy attacker *PA*'s point of view. Naturally, it follows that the larger the AS, the better the privacy. This also means that in a given vehicle traffic situation, heavy vehicle traffic situation tend to ensure more privacy due to the number of vehicles participating in communication and pseudonym change [16,19]. The entire AS in our case would comprise the set of all $\mathcal{V}$. However, it is not possible to have the entire vehicle population as the AS since our network is disconnected as we shall explain later.

Tracking Probability. The tracking probability, T_p of the PA over a vehicle, V_i, is the probability that the anonymity of a vehicle in a density zone is equal to 1. The tracking probability, T_p can be derived as follows, suppose we have $D_Z = \{Z_1, Z_2, Z_3, \ldots Z_n\}$ density zones, where a vehicle V_i is located in a zone Z_i during a short duration of time, $t = I_A$ where I_A is the anonymity duration (described later), then the probability of tracking by the adversary within zone D_i can be expressed as

$$T_p = Pr(|AS| = 1) \tag{1}$$

In practical terms, from the PA's point of view, this means that a vehicle has no anonymity when the system has $|AS| = 1$. Similarly, the composite anonymity

of the vehicles within a zone can be calculated by the number of vehicles that meets the $|AS| = 1$ criteria. A density zone where 30% of the vehicles have an $|AS| = 1$ (i.e. $T_p = 0.30$) can be said to guarantee an anonymity of 70% (i.e. 1 - $T_p = 0.70$ cannot be tracked).

Entropy. Although entropy generally means the degree of disorderliness of a system as defined in set theory. In the context of location privacy, it is a measure of anonymity according to Shanon's theory of information. Shanon's theory have been widely used in the evaluation of location privacy for vehicular networks. The uncertainty in the connection rate, the random mobility and unpredictability of our system allows us to model entropy based on Shanon's equation as [20] as follows. Let V be a discrete random variable, which is the number of vehicles, with a probability mass function $P(V = V_i)$ where $i = \{1, 2, \ldots, n\}$, then the entropy H_V of the AS can be expressed as below where p_i is the probability of each vehicle being the target of the adversary, where N represents the total number of observed vehicles by the adversary.

$$H_V = -\sum_{i=1}^{N} p_i \log_2 p_i. \tag{2}$$

Usually, the value of entropy can be normalized to have values with the domain of [0, 1]. This makes it possible to compare the entropy value with the maximum entropy of the system which is the uppermost limit of of H_V as follows

$$H_{max} = -\sum_{i=1}^{N} p_i \log_2 p_i = \log_2 |N| \quad \text{if } \forall i : p_i = \frac{1}{|N|} \tag{3}$$

The degree of anonymity is then

$$d_A = \frac{H_V}{H_{max}} \tag{4}$$

Anonymity Duration. In characterizing the anonymity of our system, we derive the anonymity duration, I_A as the time taken by a vehicle V_i to negotiate and change pseudonyms within a zone, Z_i. Note that our system is disconnected, hence we can only effectively evaluate the activity around each mix-zone in isolation. For this reason, we assume that the PA is running some tracking algorithm with which it tries identify target vehicles by matching their identities with different probability values.

There is an RSU located at every point-of-interest (PoI) location in our simulation area which are indeed the density zones described earlier. The connection rate is assumed to follow a Poisson process as in [12,15]. Let $T = I_A$ be the average time interval within which the RSU records vehicles connection activities to it. Again, let V be a random variable which is the number of vehicles that come in contact with RSU_i at density zone Z_i during I_A (i.e. during T),

V being the AS. Finally, let the inter-arrival time between connections have an exponential distribution with a mean value of $1/\lambda$. The anonymity interval is the time duration within which vehicles try to change pseudonym. It is also during this time that the PA monitors vehicles for tracking and possible identification. For a disconnected network, this time is not continuous but can be measured in snapshots of minutes or a few hours. Indeed, we regard (T as the anonymity interval I_A defined earlier), then the probability that ($V = V_i$) at ($T = I_A$) can be expressed as the Poisson process in (5).

$$P(V = v_i|T = T_A) = \frac{(\lambda t)^{v_i}}{v_i!} e^{-\lambda t} \tag{5}$$

The adversary's intention is to identify a target after pseudonym change in the mix-zone within the anonymity duration, I_A. However, not all vehicles within a density zone can successfully change pseudonyms. This can be attributed to such vehicles not being qualified enough to be considered pseudonym change candidates by other vehicles due to poor reputation records. Hence, we can define the expected anonymity set of vehicles, V_E as

$$P_{E_x}(V_E = V_i|T = I_A) = \sum_{i=1}^{\infty} v_i \frac{(\lambda I_A)^{v_i}}{v_i!} e^{-\lambda I_A} = \lambda I_A \tag{6}$$

V_E is essentially, the average number of vehicles expected to connect and disconnect with the RSU during I_A can be expressed as

$$E_X(V|T = I_A) = \sum_{i=1}^{\infty} v_i \frac{(\lambda I_A)^{v_i}}{v_i!} e^{-\lambda I_A} = \lambda I_A \tag{7}$$

4 Simulation and Anonymity Analysis

In this section, we describe our experiments and conduct performance evaluation of our scheme to understand its effectiveness. Our analytical model is supported by simulation results.

4.1 Simulation Setup

We implement our scheme using a popular and widely used network simulator for delay tolerant networks namely the *Opportunistic Networking Environment* (ONE) simulator [21]. The ONE simulator has been used to investigate several application scenarios for VANETs [5,22]. We evaluate the performance of our system under a specific use case of an intermittent/disconnected VANET. Our simulation runs involves 400 vehicles and 7 stationary relay nodes as RSUs. Table 1 presents a detailed summary of our key simulation parameters. The vehicles move on the map of the City of Helsinki which is the default map in the ONE simulator measuring 4500×3400 m^2. The RSUs are placed at chosen intersections which are the epicentres of density zones as shown in Fig. 1.

Fig. 1. Snapshot of the Helsinki city map

In accordance with Finnish traffic regulations, the average lower and upper speed bounds for vehicles is 30 to 60 kmh^{-1}. We set all vehicles in our experiments to drive at the uniform upper bound limit of 60 kmh^{-1} to ensure a uniform arrival rate at the density zones. Since vehicles usually follow defined routes in the form of roads, our model assumes each vehicle follows the *shortest path map-based movement* mobility model where vehicles are first situated randomly on different spots on road and then allowed to travel along predefined routes to their destinations. Different from our benchmark model, we deployed 400 vehicles and ran an extended simulation of 1 and 2 h respectively for values of the I_A. This is due to nature of our disconnected network environment that requires adequate number of vehicles to generate the desired statistics for analysis. We conduct our experiment only on top of the inbuilt PRoPHET routing protocol in the ONE simulator [23].

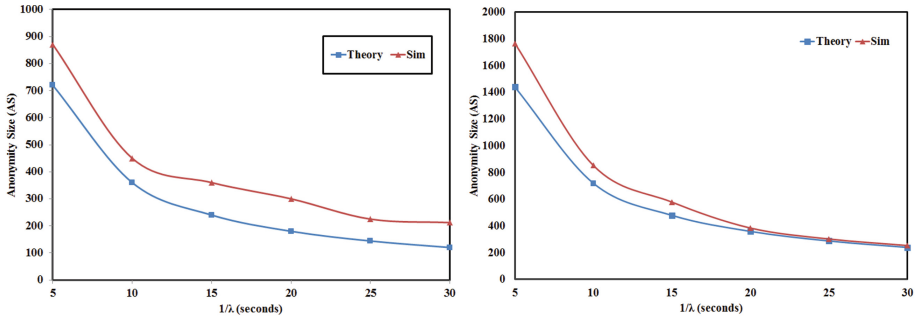
Table 1. Simulation settings

Simulation parameter	Settings/Description
Sim duration	1 & 2 h
Number of vehicles and RSUs	400 vehicles; 7 RSUs
Vehicle speed	30 kmh^{-1} – 60 kmh^{-1}
Transmission coverage	100 m
Mobility model	Shortest path map based movement
Packet size	500 k – 1 M
Message generation interval	25 s – 35 s

4.2 Analysis and Evaluation

The results from our experiments (marked Sim) are compared with our analytical model (marked Theory) as shown in Fig. 2 for different values of vehicle arrival rate, λ. We assumed a vehicle arrival of 5 vehicles per second up to 25 vehicles (i.e. $1/\lambda = \{5, 10, 15, 20, \text{and } 25\}$ with an increment of 5 and $I_A = 3600$ (1 h) and 7200 (2 h) seconds. In both Figs. 2a and b, we observe that the AS gradually depreciates as the vehicle arrival rate increases for both I_A values. This is because less frequent arrival rates means that fewer vehicles arrive at a density zone. The behavior of the graph also corroborates with the known fact that in reality, vehicles can avoid density certain zones that are notorious for low vehicular density since they have less chances of meeting pseudonym candidates in such zones compared to those known for more vehicular density. Again, we see that the values of the AS for $I_A = 7200$ is higher than that for $I_A = 3600$ which is in agreement the fact that the higher the vehicle density in a network, the higher the achievable privacy, and by extension the more chances of vehicles finding pseudonym change partners in a network. According to the work in [12], vehicles wishing to enjoy high privacy should take advantage of density zones that are notorious for high vehicle arrival rate and density to negotiate and change pseudonyms.

From existing literature, it is established that the distribution of the AS has a direct ratio to the I_A. Hence The decrease in the AS reflects the nature of our network where vehicles have temporary and intermittent connections. When compared to our baseline model [12,15] where the authors analysed pseudonym change at a small social spot, our analysis agrees with their model. A small social are temporary meeting points such as traffic intersections as against large social spots such as parking lots and shopping where vehicles meet for a longer duration of time running into hours. Note that the anonymity is based on the premise that connections can be sustained for the duration of I_A, hence the 40 s duration yields better values for the AS. This supports the notion that larger vehicle density due to a more frequent arrival rate favours a better anonymity where



(a) $I_A = 3600$ seconds (b) $I_A = 7200$ seconds

Fig. 2. Anonymity analysis for different anonymity intervals I_A

vehicles encounter more potential candidates with which to change pseudonyms at a density zone. Less frequent arrival rates means that fewer vehicles choose to use a density as pseudonym change points.

The use of live traffic information from navigation platforms such as Google Map for travel route planning is quite popular these days through which vehicle drivers can avoid certain routes based on the estimated traffic delay. Again, this suggests that higher arrival rate and more waiting time yields less anonymity. As we can see from both figures of the anonymity curve, the simulation results agree with previous work where longer (infrequent) arrival rates reduces the number of vehicles that aggregate at a zone. This leads to a situation where vehicles are discouraged from using a density that is prone to low vehicle density.

5 Conclusion and Future Work

In this paper, we added our voice to efforts to characterize and analyse privacy for a disconnected VANET using the use-case of a ventricular delay tolerant network. Our analysis follows established schemes for anonymity analysis in VANETs with varying vehicular density. We validate our analytical model with experimental results. Different from other schemes, to the best of our knowledge, this is the first attempt to analyse the privacy in a disconnected VANET. We note that our work is in progress, albeit has provided a new scope for further research in this area. In the light of this, in our future work, we are interested in developing a scheme to validate and analyse the anonymity of vehicles users using empirical data to compare with our experimental results. We also intend to develop a privacy solution more suited for disconnected network environments. Our experiments can also be performed on more robust and heterogeneous mobility scenarios such as those combining vehicles and pedestrians.

Acknowledgement. The funding for this work is from the Overseas Scholarship Scheme (OSS) of the Petroleum Technology Development Fund (PTDF) of the Federal Government of Nigeria with support from the PETRAS Project (in conjunction with IoTUK) and the Institute for Communication Systems, home of The 5G Innovation Center (5GIC), University of Surrey, Guildford, United Kingdom.

References

1. Greenblatt, N.A.: Self-driving cars and the law. IEEE Spectr. **53**(2), 46–51 (2016)
2. Ryan, P.: The driverless truck is coming, and its going to automate millions of jobs (2016). https://techcrunch.com/2016/04/25/the-driverless-truck-is-coming-and-its-going-to-automate-millions-of-jobs/
3. Raya, M., Papadimitratos, P., Hubaux, J.-P.: Securing vehicular communications. IEEE Wirel. Commun. Mag. **13**, 8–15 (2006). Special Issue on Inter-Vehicular Communications
4. Banerjee, N., Corner, M.D., Towsley, D., Levine, B.N.: Relays, base stations, and meshes: enhancing mobile networks with infrastructure. In: Proceedings of the 14th ACM International Conference on Mobile Computing and Networking, San Francisco, California, USA, pp. 81–91. ACM, New York (2008)

5. Pereira, P.R., Casaca, A., Rodrigues, J.J.P.C., Soares, V.N.G.J., Triay, J., Cervello-Pastor, C.: From delay-tolerant networks to vehicular delay-tolerant networks. IEEE Commun. Surv. Tutorials **14**(4), 1166–1182 (2012). IEEE Press, New York
6. Mohamed, N.M., Jalel, B.-O., Mohamed, H.: Survey on VANET security challenges and possible cryptographic solutions. Veh. Commun. **1**(2), 53–66 (2014)
7. Freudiger, J., Raya, M., Félegyházi, M., Papadimitratos, P.: Mix-Zones for location privacy in vehicular networks. In: Proceedings of the First International Workshop on Wireless Networking for Intelligent Transportation Systems (Win-ITS) (2007)
8. Verma, M., Dijiang, H.: SeGCom: secure group communication in VANETs. In: 6th IEEE Consumer Communications and Networking Conference (CCNC), pp. 1–5 (2009)
9. Beresford, A.R., Stajano, F.: Location privacy in pervasive computing. IEEE Pervasive Comput. **2**(1), 46–55 (2003). IEEE Educational Activities Department, Piscataway, NJ, USA, ISSN: 1536-1268
10. Leping, H., Matsuura, K., Yamane, H., Sezaki, K.: Enhancing wireless location privacy using silent period. In: IEEE Wireless Communications and Networking Conference (WCNC), vol. 2, pp. 1187–1192 (2005). ISSN: 1525-3511
11. Suguo, D., Haojin, Z., Xiaolong, L., Ota, K., Mianxiong, D.: MixZone in motion: achieving dynamically cooperative location privacy protection in delay-tolerant networks. IEEE Trans. Veh. Technol. **62**(9), 4565–4575 (2013). ISSN
12. Rongxing, L., Xiaodong, L., Luan, T.H., Xiaohui, L., Xuemin, S.: Pseudonym changing at social spots an effective strategy for location privacy in VANETs. IEEE Trans. Veh. Technol. **61**(1), 86–96 (2012). ISSN: 0018-9545
13. Chibueze, P.A.O., Haitham, C., Zhili, G., Ganesh, C., Yue, C., Philip, M.A., Masoud, A.T.: Privacy-enhanced group communication for vehicular delay tolerant networks. In: 9th International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST), pp. 193–198 (2015)
14. Soares, V.N.G.J., Farahmand, F., Rodrigues, J.J.P.C.: A layered architecture for vehicular delay-tolerant networks. In: IEEE Symposium on Computers and Communications (ISCC), pp. 122–127 (2009)
15. Rongxing, L., Xiaodong, L., Luan, T.H., Xiaohui, L., Xuemin, S.: Anonymity analysis on social spot based pseudonym changing for location privacy in VANETs. In: IEEE International Conference on Communications (ICC), pp. 1–5 (2011). ISSN: 1550-3607
16. Tomandl, A., Scheuer, F., Federrath, H.: Simulation-based evaluation of techniques for privacy protection in VANETs. In: IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), pp. 165–172 (2012). ISSN: 2160-4886
17. George, P.C., Huirong, F., Abdelnasser, B.: Evaluating location privacy in vehicular communications and applications. IEEE Trans. Intel. Transp. Syst. **9**(17) (2016). ISSN: 2658-2667
18. IEEE Standard for Wireless Access in Vehicular Environments Security Services for Applications, Management Messages. In: IEEE Std 1609.2-2013 (Revision of IEEE Std 1609.2-2006), pp. 1–289 (2013)
19. Hassan, A., Noor, A.: A pseudonym management system to achieve anonymity in vehicular ad hoc networks. IEEE Trans. Dependable Secure Comput. **13**, 106–119 (2016). ISSN: 1545-5971
20. Claude, E.S.: A mathematical theory of communication. Bell Syst. Tech. J. **27**(3), 379–423 (1948). ISSN: 0005-8580

21. Ari, K., Jörg, O., Teemu, K.: The ONE simulator for DTN protocol evaluation. In: SIMUTools 2009: Proceedings of the 2nd International Conference on Simulation Tools and Techniques, New York, NY, USA. ICST, Rome (2009). ISBN: 978-963-9799-45-5
22. Rongxing, L., Xiaodong, L., Xuemin, S.: SPRING: A social-based privacy-preserving packet forwarding protocol for vehicular delay tolerant networks. In: Proceedings of the 29th Conference on Information Communications, pp. 1–9 (2010). ISSN: 0743-166X
23. Lindgren, A., Doria, A., Schelén, O.: Probabilistic routing in intermittently connected networks. SIGMOBILE Mob. Comput. Commun. Rev. **7**(3), 19–20 (2003). New York, USA, ISSN: 1559-1662