

Utilizing Social Links for Location Privacy In Opportunistic Delay-Tolerant Networks

Sameh Zakhary and Milena Radenkovic

School of Computer Science
University of Nottingham,
Nottingham, NG8 1BB, UK
Email: {szz, mvr}@cs.nott.ac.uk

Abstract—This paper is concerned with improving location-privacy for users accessing location-based services in opportunistic DTNs. We design a protocol that offers location privacy through request/reply location obfuscation technique that uses the nodes' own social network to drive the forwarding heuristic. We propose a fully distributed social-based location privacy protocol (*SLPD*) that utilizes social ties between nodes to ensure K-Anonymity, i.e. the requesting node's locations cannot be determined from at least k-1 other nodes in its social network. We evaluate *SLPD* using extensive simulations and real connectivity data traces. We compare our results to a benchmark protocol that requires centralized trusted server. We show that our distributed protocol is applicable to DTNs with various mobility patterns, and provides the user with the required privacy at less than 30% of the privacy range we define. *SLPD* achieves success ratios similar to the ones obtained using centralized benchmark solutions up to 15% privacy requirements.

Keywords- location privacy; location-based services; delay-tolerant network; opportunistic network; Security

I. INTRODUCTION

Privacy and security are becoming a major concern in DTNs. Recently there are increased number of services that enable monitoring and tracking of human activities [1]. For example, selling/auction sites collect personal information about the users' buying habits and activities to provide tailored services [1]; Location Service Providers (LSPs) collect information about users' locations [2]. Governments also support gathering personal data by justifying the need to track and identify people for safety and security reasons, or by setting legislation support for data sharing to enable massive data collection [3]. Recent examples of data sharing between entities without user's consent and sensitive data leakages have raised the public alertness to how their data can be used or taken advantage of. Due to the wide deployment of mobile devices capable of networking with each other and accurately sensing user's location [4], location privacy has become a greater concern than in traditional networks and it is particularly difficult to achieve a satisfactory level of location privacy in situations where nodes rely on location-based services (LBS).

Many networks paradigms suffer from location privacy issues. For example, in Pocket-Switched Networks (PSNs) [5], tracking of mobile devices' locations at different times is equivalent to tracking their owners. In Vehicular Ad-Hoc Networks (VANET), privacy is increasingly becoming a major issue [6] due to the wide deployment and the easiness of tracking. In these networks, a privacy-concerned user may avoid participating by disabling their mobile device's opportunistic-networking capability at different times, hence causing the network to become more fragmented. We are particularly concerned with the impact of location privacy on

the feasibility of communication in an opportunistic network. Some recent work [7] shows that users' location privacy preferences may have a huge impact on the delivery ratio in such networks i.e. privacy concerns could eventually lead to a completely disabled communication.

Our contributions are to provide a discussion and evaluation of the impact of location-privacy in a real-life social and heterogeneous city scenarios, as well as to design, implement and evaluate a hybrid location privacy-preserving protocol for opportunistic DTNs. We propose *SLPD* which is a hybrid protocol that combines social-based forwarding [8] for maintaining location privacy stage, label-based forwarding [9] for addressing the LBS-reply back to the requester, and DTN forwarding as the underlying protocol. We design *SLPD* for accessing location-based services in networks similar to [10]. For evaluation, we use extensive simulation using real data traces and map-based application scenarios. Our results show that a fully distributed socially driven location privacy protocol is feasible for DTNs at low privacy requirements.

The rest of the paper is organized as follows. Section II gives an overview of the related work. Section III presents our proposed social-based location privacy protocol, as well as implementation and design details. Section IV details the evaluation methodology and results. Section V concludes the paper and looks into future work.

II. RELATED WORK

Location privacy has proven to be difficult to define [11], achieve, or even support at a satisfactory level in DTNs. Recent research has focused on assessing the impact of privacy-concerns on the DTN performance [7]. Especially, when nodes rely on LBS for the user applications to function properly, and where the personal identity of the user is attached to this node. For example, in Pocket-switched Networks [5], a mobile phone carried by a person is associated with that person identity and the detection of this mobile phone location at different times is almost equivalent to tracking its owner. Parris et al. [12] propose using obfuscation of user's own social network exchanged as part of the social-forwarding protocol in DTN, by altering or hiding the social network information, in order to prevent the attacker from inferring any physical proximity between nodes on the message delivery path. These techniques affect forwarding but are not applicable when the node has to access LBS and provide location information for the request to be answered correctly.

In VANET, Rongxing et al. [13] propose using static road side units that are strategically positioned at high-social intersections to temporary store packets for passing vehicles to provide additional privacy. This work assumes that RSUs are trustable and non-compromisable, which is specific to

small deployments where all RSUs are under a single management authority and unsuitable for many real-world scenarios. In addition, RSUs become a high-risk to the entire network if any RSU gets compromised, and they become prominent targets for malicious attack to compromise the whole network and collect private information about all vehicles in VANET.

Other applications in DTNs are even more sensitive to location privacy, such as in the military domain, where an enemy could destroy key nodes –for example, example command unit- if their location can be tracked through traffic back tracing or flow analysis. Recent research efforts have been looking at the problem of information publishers’ or subscriber’s anonymity where publishers/subscribers require to be anonymous and hide their identity from each other and from any other node in the network [14]. Such problem is often referred as source unobservability in the case of the publishers [15]. Most of the solutions proposed assume either dense networks or trusted servers, while other solutions rely on multiple trusted nodes injecting dummy packets into the networks. These dummy packets serve the purpose of subverting the attacker’s efforts and making them unable to identify interesting packets – that they could subsequently trace back to its source. If this dummy traffic is not monitored and controlled, it causes increased un-useful traffic that results in what is known as “Traffic Explosion” that negatively affects useful network traffic and causes data delivery ratios to drop sharply. Unhandled or malicious traffic explosion situation can further lead to a complete denial of service in the network [16]. Recently, network coding with encryption has been proposed as a convenient way to absorb these dummy packets and eliminate any reliance on trusted party in the network [15], but it suffers from high computational and deployment overheads. Pongaliur and Xiao [17] propose a protocol to maintain source privacy. The protocol modifies the messages dynamically by dynamically selected nodes to make it difficult for a malicious entity to trace back the packet to a source node. This protocol requires distribution and management of cryptographic keying materials between all nodes, and this is not applicable to most opportunistic DTNs.

III. EXTENDING K-ANONYMITY WITH SOCIAL LINKS IN DTNS (SLPD)

A. Overview

Our approach is to utilize social-links (i.e. friendship relationship between involved users in opportunistic DTNs) to maintain the user’s privacy while accessing location-based services, and attempt to provide the user with location privacy based on K-Anonymity technique. We refer to this approach as “Social-based Location Privacy in DTNs” (or **SLPD**). We structure the user *social profile* as an n-tuple *profile attributes*, and each of these attributes is assigned a different weight that reflects the relevance of this attribute to the whole profile. E.g. affiliation attribute could be given 40% importance, where the gender could be given 5% importance in matching the two nodes’ profiles.

Since DTNs are mostly disconnected, and contacts are short lived with no end-to-end connection, users need to utilize their encounters as well as leverage the social relations efficiently with other nodes. We build on the fact that users trust their contacts –friends and relatives in their social network- with providing location obfuscation when an encounter happens, rather than trusting a completely unknown user(s) which they encounter randomly in an opportunistic DTN. Hence, we assume a trusted-community

security model. Only external attackers can be capable of eavesdropping on some of the traffic in the network.

A user who wishes to request LBS searches for nearby friends, and forwards a copy of their request to one of the available friends. Their friend then forwards the request intra-social group using social forwarding for a defined number of hops (K) within the user’s social network based on predefined social profile match criteria. After the first K hops, the message can be forwarded using any DTN forwarding protocol (*SnW* in our evaluation case). This allows the requester not to reveal its location by directly contacting the LBS, but relay on the social forwarding protocol to form an obfuscation path to the LBS. Similar to K-Anonymity, the LSP will not be able to distinguish the origin of the location request from at least K other users who have participated in forwarding the message. Moreover, since socially related users would normally be co-located and exchange other types of messages, there is little benefit for the attacker to carry traffic analysis attack on the network. This approach is different from Onion Routing [18] in that the forwarding nodes are not connected end-to-end between the requester and the LSP, but the forwarded using DTN protocol.

B. Design and Implementation

We build a social layer that controls how a DTN forwarding protocol chooses the possible next hop(s). Without loss in generality, we implement this layer on top of one of the existing quota-based DTN forwarding protocol, namely “*Spray And Wait*” (*SnW*) [19]. This social layer is separated from the forwarding layer as it examines social properties (profile attributes and social links) and influences the forwarding decision based on the topology status as sensed from the wireless medium. We simplify this approach using figure 1, where we show how the social-based obfuscation is performed to guide the DTN forwarding protocol. The social layer is responsible – only in the first $k-1$ forwarding decisions - for searching and providing a list of neighbouring nodes that match a given social criteria to maintain the requesting node’s privacy. This list is then used by the forwarding protocol, and a number of these possible nodes are forwarded a copy of the message (e.g. for *SnW* depending on the remaining number of message copies at the current node, only a subset of these possible next hops gets a copy of the message until no copies are left at the current node for the first $k-1$ hops).

We propose that each node in the social group which is on the obfuscation path maintains a mapping table of which requests that it has forwarded belong to which requesting nodes. This aims to speed-up the LBS-reply addressing. Each of the $k-1$ nodes does the reverse operation as soon as a message arrives at any of the $k-1$ nodes originally participated in the obfuscation forwarding of the corresponding request. This is similar to the addressing used in Gently [20] where context awareness and group labelling are used to determine the opportunistic forwarding decision. Our approach differs in that the underlying forwarding protocol is used once the reply hit any of the nodes capable of re-addressing (i.e. setting the address of the original requester instead of the group label).

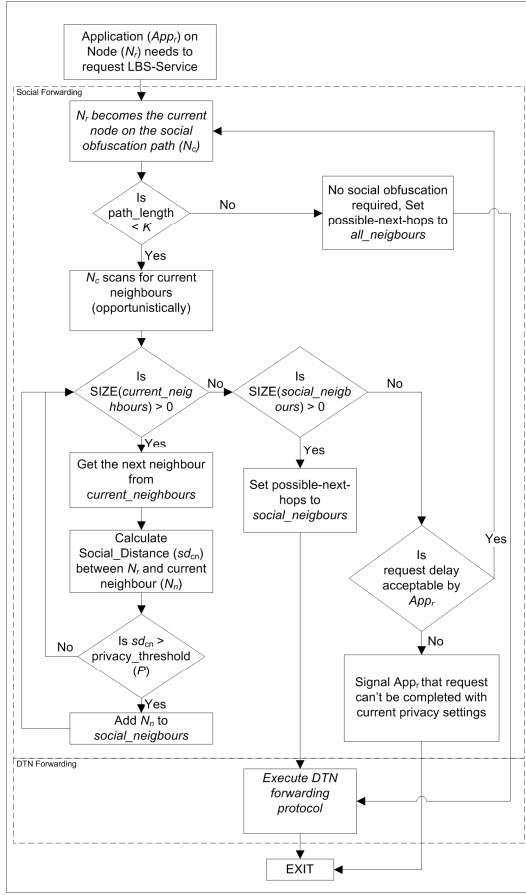


Figure 1. Social-based location privacy in Opportunistic and social DTNs (SLPD). Social path obfuscation protocol for K-Anonymity.

Due to the asynchronous operation of the forwarding in DTN, It is possible that the LBS-reply never reaches any of the nodes on the original obfuscation path on its way back from the LBS, i.e. none of the $k-1$ intermediaries are on the return path. However, in this case, the intermediaries are still able to forward the reply using the group label similar to [20] in an attempt to reach the original requester without knowing its exact identity. Intermediary nodes follow the same process in figure 1 for $k-1$ times as instructed by the requesting node (N_r). Since the social layer guides the forwarding protocol on which nodes to select from the current neighbouring nodes, then there is at least $k-1$ other socially matching nodes (through calculating sd as it will be shown later) similar to the sender, where each has participated in store-carry-forward of the LBS-request. The outbound node of the intermediate nodes - last node number $k-1$ on the obfuscation path - removes the identity of the original requester (N_r) and replaces it with the social group label, similar to [9], before forwarding the message freely to any neighbouring node (i.e. not following the social criteria).

The sender of a message calculates the Social Distance (sd_m), which is defined as the weighted number of matching *profile attributes* between N_r and its current neighbors N_n , and it is calculated as in (1).

$$sd_m = M(P_r, P_n) / C(P_r) \quad (1)$$

Where function $M()$ searches for the matching degree between the social attributes of the two given social profiles, $C()$ returns the number of attributes in the given profile, and P_n represents the social profile of a node (n). Matching

degree reflects both the number of matching *profile attributes* as well as the weight each of these attributes is given in the *social profile*.

IV. EVALUATION METHODOLOGY AND RESULTS

A. Oracle-Based K-Anonymity (OBK) – Location Privacy Preserving Forwarding - Protocol

We first implement a benchmark k-anonymity protocol. This is in order to evaluate the effectiveness of K-Anonymity using a centralized matchmaker server as proposed in [21], but in DTN scenario. K-Anonymity is a very popular approach for location anonymity in traditional and P2P networks, where user's location is indistinguishable from locations of at least $k-1$ other users. We propose a privacy-preserving protocol for accessing location-based services, which is based on the concept of K-Anonymity and extends it to utilize social links between nodes in the social DTN mobile nodes. Next we show a pseudo code in algorithm I, similar to [21], with critical extensions required to make it applicable to DTNs. In [21], it is assumed that a centralized and trusted matchmaker server is always accessible to the nodes, i.e. nodes have access to that server to send their query and this server forward an anonymized request to the LSP. However, for our benchmark protocol, we propose to limit the role of the matchmaker server to only provide a list of candidate nodes that match the profile criteria set by the requester. This list contains information about nearby nodes that are willing and within the cloaked range to provide K-Anonymity to the requesting node. Contrary to [21], we extend the requester capability so that it carries the forwarding of its anonymized LBS request using spray-and-wait through the provided set of matching nodes as intermediaries to the LSP. We refer to this protocol as **Oracle-Based K-Anonymity (OBA)**, and the Oracle is represented by the centralized-trusted matchmaker. **OBA** serves as a benchmark for the possible success ratio (*SR*) and privacy combination in a given scenario.

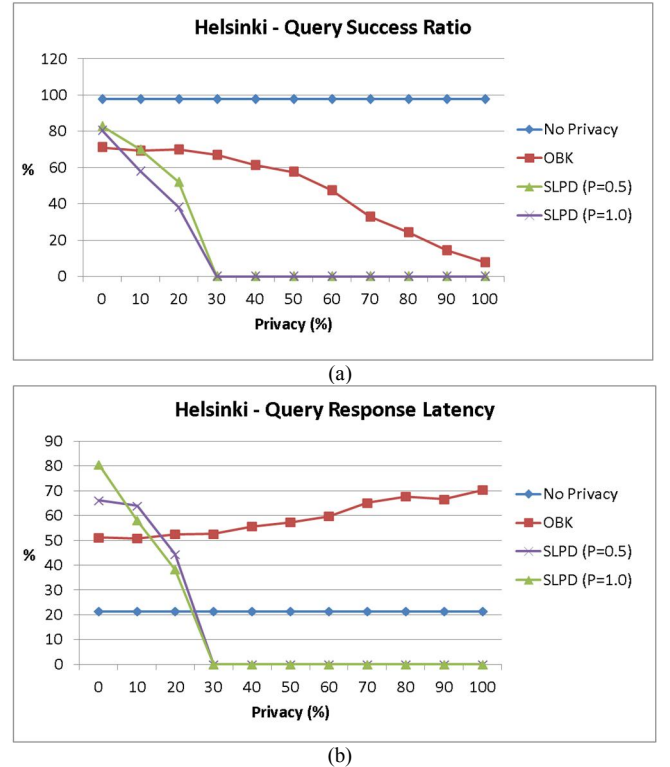


Figure 2. Helsinki city-scenario - Comparing LBS query (a) success ratio and (b) round-trip latency across different protocols.

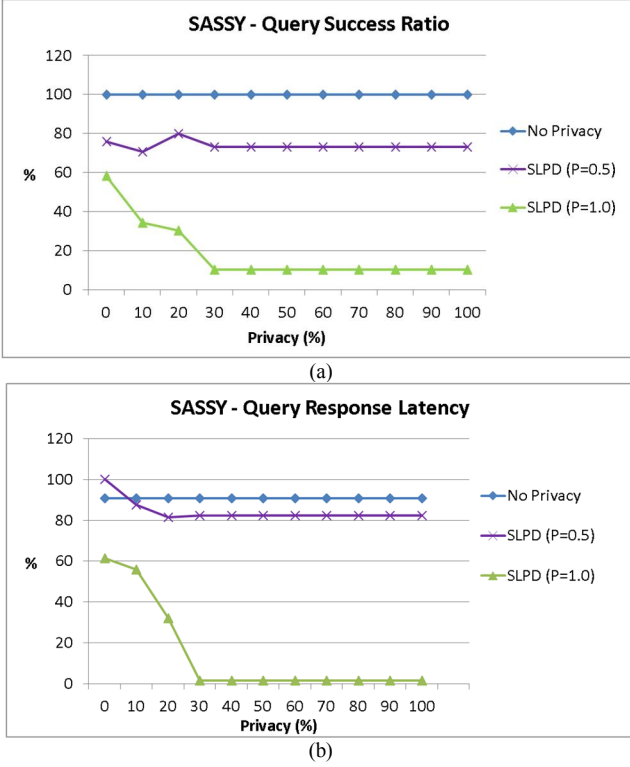


Figure 3. Real-world traces scenario- Comparing LBS query (a) success ratio and (b) round-trip latency across different protocols.

B. Experimental Results

1) In the Heterogeneous City-Center Scenarios

We implement our protocols in Opportunistic Network Environment (ONE) simulator [22]. We use “Map-Based Movement” models. These models utilize map data in order to constrain node movement to the streets, roads and tram path according to the node type (Pedestrians, Cars or Trams). Each point in the figures is the average of 20 runs around the map (*Helsinki City Map*).

We evaluate our social-based anonymity *SLPD*, and compare it to the benchmark *OBK* under varying degree of privacy requirements. The privacy requirements are reflected on the user choice of the value of K and P , i.e. the greater this value, the more it is harder to determine the user and hence stronger privacy. K values were varied between 3 and 13 for both protocols, and results have been recorded. *SLPD* was evaluated using varying social matching threshold, i.e. users can select any value for P , where it refers to the “social distance” which measures the level of matching in social links between the nodes when they encounter each other. Users who desire more privacy set P to a value close to 1, and vice versa. We assume that location service provider (LPS) is connected to the infrastructure and at the same time can communicate opportunistically using Bluetooth. Location queries are sent every 500 seconds from one group of 20 tourist walking and querying locations around the city. Each member of this group has a defined social profile (e.g. interest and friends), and it is used to identify other nodes that are either friends or that has common social links/common interests.

Figure 2.a presents *SR* of our Social-based privacy protocol compared to two different bench-marks (i.e. no privacy social forwarding and centralized oracle-based approach at low privacy requirements (15)). The query *SR* was recorded as a ratio of the maximum achievable result in each scenario. The results show that the maximum

achievable *SR* is close to the oracle-based approach at lower privacy requirements, but falls sharply with increasing privacy or profile matching P . This is due to the lost communication opportunities with other nodes that do not belong to the sender’s or intermediaries’ social group. It is interesting to examine the social ranking techniques to allow the protocol to utilize the social overlay network similar to the one proposed in [23]. We can see that for our city center scenarios, the users are able to communicate while maintaining up to 30% privacy ($K \leq 6$) without the need for centralized servers, but failed at higher privacy requirements.

Figure 2.b shows the latency of the query response, which is measured by the round-trip time as a percentage to the worst latency in the corresponding scenario, i.e. the delay between the users sending the LBS-request until the reply is received. We can observe that *SLPD*’s latency drop sharply as the privacy requirements increase and in line with the drop in *SR*. Interestingly, as P decreases *SLPD* show higher *SR* and less-sharp drop as compared to in case where ($P=1$). On the other hand, *OBK* shows a steady increase in latency, and this is despite the fall in *SR*; which indicates that even with a centralized server for match-making, the user will still suffer from lower success ratio and higher latency.

2) Real-World Connectivity Data Traces with Social-Links

We use real-world connectivity traces that record the encounters between different participants throughout the monitoring duration, and we also utilized the social information between these participants so that the profile matching can be derived from the social-graph information. We use real-world SASSY [12] dataset contributed by Bigwood et al., where they have recorded the mobile encounters of 27 participants (22 undergraduate students, 3 post-graduate students and 2 staff members) carrying T-mote sensor nodes in their day-to-day life activities over the period of three months. The range of these devices was about 10m, and encounters were uploaded to a base-station regularly. Researchers have also collected the social-network relation (i.e. friendship relation from Facebook) between these participants to represent the pre-existing social network, or Self-Reported Social Networks (SRSNs) [24].

We have experimented with the complete dataset, but omit the details due to space, and show evaluation using two weeks period. We calculated the social-profile matching probability P as follows: 1) direct friends \rightarrow ($P=1$), 2) Friends-of-friends (or second level friends) \rightarrow ($P=0.70$), 3) Third level friends \rightarrow ($P=0.30$), and any further levels are considered to have ($P=0.0$) (i.e. no matching profiles attributes is assumed to exist with individuals who are more than three levels down the social graph).

Figure 3.a. shows that *SLPD* could achieve higher and more steady *SR* for moderate-privacy requirements ($P=0.5$), in real-world traces scenario, compared to the “No privacy” scenario. While for the high-privacy requirements ($P=1$), the *SR* is sharply declining as K increases. The results for ($P=0.5$) is better due to the reliance on additional nodes (or friends-of-friends) to provide obfuscation, while for ($P=1$) only direct friends could be trusted to provide obfuscation. A considerable tradeoff between privacy requirement, user’s own social network size and mobility pattern can be seen using this real-world dataset, which have contributed to the overall better results compared to the *city-center* scenario. Users were able to achieve higher *SR* relying on nodes in their own social network that is also co-location or encounter

N_r = the original LBS requester node.
 K = the minimum number of nodes the requester want to be indistinguishable from.
 P = the profile attributes of a node.
 $P_{criteria}$ = the profile attributes for an anonymity request (i.e. female and/or walking users).
 S_c = anonymity-square area centre
 l = anonymity-square side length in meters

- 1: For each node (N_i): node sends a report of its location and profile attributes to a trusted centralized server (Loc_{cent}), and follows with subsequent location updates when it changes location.
- 2: The requester (N_r) locates the nearest LBS server to its current location (LBS_{srv}), so that it can be contacted with a specific query.
- 3: User of node (N_r) sets the anonymity requirement of the LBS application (App_r) running on the node in terms of (S_c , l_r , $P_{criteria}$).
- 4: N_r determines its current location as (x_r , y_r) through built-in sensor.
- 5: Node (N_r) sends a query Loc_{cent} to obtain the number of nodes available within the cloaking region defined by (S_c , l_r).
- 6: Loc_{cent} searches its database for all possible nodes that matches the criteria, and replies to N_r with M_{match} that represents the number of nodes matching the anonymity criteria.
- 7: **IF** ($M_{match} \geq K$) **THEN**
- 8: App_r is signaled to generate an LBS-request. The LBS-request is then forward to surrounding nodes to attempt to deliver it to LBS_{srv} .
- 9: **ELSE**
- 10: The user is notified that App_r can not send the request and maintain the user pre-set location privacy requirements.
- 11: App_r halts awaiting signal to start transmitting, then **GOTO** 4
- 12: **ENDIF**
- 13: LBS_{srv} receives the LBS-request which have a location query in (S_c , l_r) and hence unable to distinguish N_r from at least K_r other nodes.
- 14: LBS_{srv} generates and forwards the LBS-reply back to N_r through social forwarding.

traces (i.e. two people who are friends are more likely to co-locate or meet in one place than two randomly chosen people).

Figure 3.b shows that *SLPD* ($P=0.5$) had higher latency, but almost always below the “No privacy” scenario, which is due to delivery more messages as shown in the SR figure. While *SLPD* ($P=1$) had lower latency that drop sharply until $K=30\%$ as that is the point where the SR drops to around 10% only.

V. CONCLUSION AND FUTURE WORK

We motivated the need for adequate protection protocols to adapt to the unique characteristics of DTNs, such as long delays and sparse neighborhood. We proposed our privacy-preserving protocol that is fully distributed and socially driven. The protocol utilized social links between involved nodes to form an obfuscation path between the requester and the LBS, so that it guarantees location privacy similar to k-anonymity for the requesting node. We showed evaluation results using two distinctive application scenarios in both city-center and social networks. And we conducted extensive simulations in two mobility models, one in the city-center Scenario using map-based movement models, and the other using real-world traces with social network (Facebook friendship). The results showed that it is possible to build social-based fully distributed and collaborative privacy-preserving protocol for accessing LBSs in DTNs. As a future work, we will evaluate *SLPD* with additional different datasets that include location information, and experiment with different user privacy preferences. We plan to expand the testing with additional simulation's movement models.

REFERENCES

- [1] Newman, A., *Protectors of privacy: regulating personal data in the global economy* 2008: Cornell University Press.
- [2] Puttaswamy, K.P.N. and B.Y. Zhao, *Preserving privacy in location-based mobile social applications*, in *HotMobile* 2010, ACM. p. 1-6.
- [3] Solove, D.J., *The digital person: Technology and privacy in the information age*. Vol. 1. 2004: NYU Press.
- [4] Pelusi, L., A. Passarella, and M. Conti, *Opportunistic networking: data forwarding in disconnected mobile ad hoc networks*. Communications Magazine, IEEE, 2006. **44**(11): p. 134-141.
- [5] Hui, P., et al., *Pocket switched networks and human mobility in conference environments*, in *SIGCOMM workshop on DTN* 2005, ACM.
- [6] Ma, D. and G. Tsudik, *Security and privacy in emerging wireless networks*. Wireless Comm., 2010. **17**(5): p. 12-21.
- [7] Parris, I. and T. Henderson, *The impact of location privacy on opportunistic networks*, in *WoWMoM* 2011.
- [8] Daly, E. and M. Haahr. *Social network analysis for routing in disconnected delay-tolerant manets*. in *MobiHoc '07*. 2007. ACM.
- [9] Hui, P. and J. Crowcroft, *How Small Labels Create Big Improvements*, in *PERCOMW '07* 2007, IEEE p. 65-70.
- [10] Mashhadi, A., B. Mokhtar, and L. Capra. *Habit: Leveraging human mobility and social network for efficient content dissemination in Delay Tolerant Networks*. in *WoWMoM*. 2009. IEEE.
- [11] Duckham, M., *Moving forward: location privacy and location awareness*, in *SPRINGL2010*, ACM: San Jose, California. p. 1-3.
- [12] Parris, I., G. Bigwood, and T. Henderson. *Privacy-enhanced social network routing in opportunistic networks*. in *PERCOM Workshops*. 2010. IEEE.
- [13] Rongxing, L., L. Xiaodong, and S. Xuemin. *SPRING: A Social-based Privacy-preserving Packet Forwarding Protocol for Vehicular Delay Tolerant Networks*. in *INFOCOM*. 2010. IEEE.
- [14] Shikfa, A., M. Önen, and R. Molva, *Privacy-preserving content-based publish/subscribe networks*. SEC, 2009: p. 270-282.
- [15] Yanfei, F., et al. *Preventing Traffic Explosion and Achieving Source Unobservability in Multi-Hop Wireless Networks Using Network Coding*. in *GLOBECOM*. 2010.
- [16] Yang, Y., et al. *Towards event source unobservability with minimum network traffic in sensor networks*. in *WiSec*. 2008. ACM.
- [17] Pongaliur, K. and L. Xiao, *Maintaining Source Privacy under Eavesdropping and Node Compromise Attacks*, in *INFOCOM2011*: IEEE.
- [18] Reed, M.G., P.F. Syverson, and D.M. Goldschlag, *Anonymous connections and onion routing*. IEEE Journal on Selected Areas in Communications, 1998. **16**(4): p. 482-494.
- [19] Spyropoulos, T., K. Psounis, and C.S. Raghavendra, *Spray and wait: an efficient routing scheme for intermittently connected mobile networks*, in *SIGCOMM workshop on DTN*. 2005, ACM.
- [20] Musolesi, M., et al., *Writing on the clean slate: Implementing a socially-aware protocol in Haggie*, in *WoWMoM* 2008, IEEE. p. 1-6.
- [21] Mano, M. and Y. Ishikawa, *Anonymizing user location and profile information for privacy-aware mobile services*, in *LBSN2010*, ACM. p. 68-75.
- [22] Keränen, A., J. Ott, and T. Kärkkäinen. *The ONE Simulator for DTN Protocol Evaluation*. in *SIMUTools*. 2009. ICST.
- [23] Dell'Amico, M. and L. Capra, *Sofia: Social filtering for robust recommendations*. Trust Management II, 2008: p. 135-150.
- [24] Bigwood, G., et al., *Exploiting Self-Reported Social Networks for Routing in Ubiquitous Computing Environments*, in *WIMOB* 2008.