

Facebook or Fakebook? The Effects of Simulated Mobile Applications on Simulated Mobile Networks

Iain Parris^{1,*}, Fehmi Ben Abdesslem, Tristan Henderson

*School of Computer Science
University of St Andrews
St Andrews, Fife KY16 9SX, UK*

Abstract

The credibility of mobile ad hoc network simulations depends on accurate characterisations of user behaviour, e.g., mobility and application usage. If simulated nodes communicate at different rates to real nodes, or move in an unrealistic fashion, this may have a large impact on the network protocols being simulated and tested.

Many future mobile network protocols, however, may also depend on future mobile applications. Different applications may be used at different rates or in different manners. But how can we determine realistic user behaviour for such applications that do not yet exist? One common solution is again simulation, but this time simulation of these future applications.

This paper examines differences in user behaviour between a real and simulated mobile social networking application through a user study ($n=80$). We show that there are distinct differences in privacy behaviour between the real and simulated groups. We then simulate a mobile opportunistic network application using two real-world traces to demonstrate the impact of using real and simulated applications. We find large differences between using real and synthetic models of privacy behaviour, but smaller differences between models derived from the real and simulated applications.

*Corresponding author.

Email addresses: isp3@st-andrews.ac.uk (Iain Parris), fba@st-andrews.ac.uk (Fehmi Ben Abdesslem), tnhh@st-andrews.ac.uk (Tristan Henderson)

URL: <http://www.cs.st-andrews.ac.uk/~ip/> (Iain Parris),
<http://www.cs.st-andrews.ac.uk/~fehmi/> (Fehmi Ben Abdesslem),
<http://www.cs.st-andrews.ac.uk/~tristan/> (Tristan Henderson)

¹Phone: +44 1334 463253

Fax: +44 1334 463278

Keywords: simulation, social networks, opportunistic networks, privacy, social-network routing
2010 MSC: 68M12, 68U20

1. Introduction

Simulation is vital for the research and development of future mobile network protocols. It can be impractical to deploy new protocols and systems on a large-scale basis under realistic conditions, especially if constant refinement and development of the systems are required. But the credibility of mobile network simulations depend on accurate characterisations of user behaviour, such as mobility and application usage. If the users being simulated use applications at a different rate to real users, or move in an unrealistic fashion, this can have an impact on the network protocols being simulated and tested.

Many of the future mobile network protocols being researched and developed, however, may also depend on future mobile applications. For instance, a disruption-tolerant mobile ad hoc network protocol might be used by future disruption-tolerant applications. How can we determine realistic user behaviour for such applications that do not yet exist? One solution is to measure user behaviour in such applications by simulating the applications themselves. But this might have a further impact on the network simulations, if behaviour in simulated applications is not the same as that in real applications.

In previous work we have conducted preliminary investigations of the difference between real and simulated applications [1], and the impact of user behaviour on mobile network performance [2]. In this paper we extend and combine this analysis, to examine differences in user behaviour between real and simulated applications and study the impact of these differences on mobile network simulation. We conduct a user study with 80 participants, half of whom used a simulated mobile social network application, while the other half used the real mobile social network application Facebook. We show that there are distinct differences in privacy behaviour between the real and simulated groups, with users of the simulated application being more willing to share data than real users. Using the results of this user study, we then simulate a set of mobile ad hoc opportunistic network routing protocols to demonstrate the impact of using real and simulated applications. Finally, we study the effects of real and simulated privacy models by comparing a synthetic privacy model to one generated from the user study results.

This paper makes the following contributions:

- We conduct, what is to our knowledge, the first study of the impact of application-layer simulations on network-layer simulations in a mobile net-

work. Our results indicate that it may be possible to use simulated applications as a substitute for real applications.

- We demonstrate that the use of synthetic privacy models can have a large impact on mobile ad hoc network simulations.
- We demonstrate that considering user behaviour, such as privacy preferences, can have a large impact on mobile ad hoc network protocol performance.

This paper is outlined as follows. Next, we characterise some of the related work. We then describe a location-privacy user study using both real and simulated users in Sections 4 and 5. Section 6 describes a privacy model based on the results of this user study. In Section 7 we use this privacy model in a set of trace-driven simulations of a mobile opportunistic network. We then discuss the implications of our results, suggest some guidelines for future simulations and research, and conclude in Section 8.

2. Related Work

Our goal is to understand the accuracy, or the credibility, of network simulation given assumptions about real-world behaviour. Many researchers have examined simulation credibility, and a good survey of research into the credibility of wireless network simulation is provided by Andel and Yasinsac [3]. Newport et al. study assumptions about radio propagation in wireless network simulations [4] by comparing simulated and real-world wireless experiments. They demonstrate that many commonly-held assumptions, such as circular transmission range, symmetric transmission capabilities, and signal strength being proportional to distance, do not hold, and this can have an impact on any simulations that make these assumptions. Tan et. al. extend this to indoor and outdoor wireless mesh networks, finding that simulations often make inaccurate assumptions about antenna diversity and path loss [5]. Torres et. al. study IEEE 802.11e wireless networks by deploying a real testbed and comparing the results to those from the ns-2 simulator [6].

More broadly, Kurkowski et al. study 114 wireless ad hoc networking papers from the ACM MobiHoc conference that employ network simulation [7]. They find that many of the papers are not credible, in that the experiments are not repeatable, are biased, or are not rigorous or statistically sound.

The problem of credible network simulations reaches beyond mobile and wireless network simulations. Pawlikowski et al. consider over 2,200 telecommunication network research papers and focus on two particular factors that might affect credibility: pseudo-random number generation and statistical analysis of the

outputted simulation data [8]. Floyd and Paxson describe strategies for attempting to simulate the global Internet [9]. Chertov and Fahmy build a measurement-based model for Internet routers and validate this model using Cisco and Juniper routers [10].

An alternative research method to simulation is to emulate a networked environment instead, by collecting traces in one real environment, and using the traces to drive experimental runs in a different testbed environment. Kim et al. describe a system for doing so and examine the benefits of using emulation [11].

In addition to understanding how physical and network layer assumptions may impact the credibility of a simulation experiment, it is also necessary to understand the effect of simulating a user-facing application. Tsai et al. [12] examine the effect of feedback in a real (i.e., non-simulated) location-sharing application tied to the Facebook social networking site (SNS).² Feedback — in the form of a list of viewers of who had viewed each published location — was found to influence disclosure choices. Although they do not investigate a simulated application, the fact that real feedback has an effect may mean that simulated feedback (e.g., using a randomly-generated list of viewers) may affect behaviour in a different way.

Consolvo et al. [13] investigate privacy concerns in a simulated social location-tracking application, employing the Experience Sampling Method to query participants *in situ* [14]. They also note problems with simulation in the post-experiment debriefings: unrealistic, “out-of-character” simulated location requests were rejected by at least one participant.

Stutzman and Kramer-Duffield [15] examine real-world privacy settings of Facebook users employing the site’s relatively restrictive “friends-only” option. More generally, Young and Quan-Haase [16] present Facebook as a case study for privacy concerns on online social network sites. Toch et al. [17] develop empirical models of privacy from a user study, as we do here, but do not apply this model to simulations as in our work.

Another factor in the credibility of a social network simulation is the realism of the social network itself. Orman and Labatut look at a preferential-attachment-based method for generating “realistic” networks, and study the impact of realism on community detection in social networks [18].

While numerous researchers have studied credibility of network simulations, and have leveraged simulated applications for user studies, to the best of our knowledge no-one has studied the impact on network simulation credibility of application simulation. In this paper we attempt to study this problem in a mobile opportunistic ad hoc networking scenario.

²<http://www.facebook.com/>

3. Privacy threats in mobile opportunistic networks

To investigate the impact of a simulated network application on network simulations, we choose to study user behaviour in a mobile opportunistic social networking application. The proposed application acts as other common online social networks such as Facebook or Twitter³, where users share information with other members of their social network, but information is transmitted over an opportunistic network, as has been proposed in such research applications as Mobi-Clique [19]. In order for such information to be transmitted, users must actively cooperate and share information with other members of the network. With this in mind, the particular user behaviour that we choose to examine relates to user privacy; in particular, users' willingness to disclose their location, which may be private, to other users: friends, family or the public at large. To be able to model and study this behaviour, it is necessary first to understand the privacy threats in such an environment, to measure users' privacy concerns and to then model these concerns.

3.1. Opportunistic networks

Opportunistic networks [20] form a popular area of current research in mobile ad hoc networking. An opportunistic network can be viewed as a mobile ad hoc network that is completely disconnected, where nodes can send messages in the absence of any knowledge of network topology. Nodes opportunistically exploit any nodes that they encounter to help messages reach their destinations. For example, users with mobile phones who encounter other users with mobile phones can exchange messages using Bluetooth, and messages could subsequently be forwarded using other encounters or other available network interfaces such as 3G. Such networks can be used to enable new applications, e.g., for social media or information dissemination [21], in the absence of any network infrastructure. In our simulated application, an opportunistic network would allow users to participate and generate or share content in an online social network even when accessing an infrastructure network is impractical, e.g., due to cost, political interference or lack of coverage.

3.2. Privacy threats in an opportunistic network

Forwarding data in an opportunistic network relies on participation by nodes in the network to route messages for each other. But if all such nodes are expected to act as routers, and some of these nodes may be untrusted, many possible privacy threats are introduced that may not exist in a traditional infrastructure network. For instance:

³<http://www.twitter.com/>

- Message contents may be disclosed, enabling them to be read by unintended parties. This threat may be mitigated with encryption, if the sender and destination are able to agree on encryption keys [22].
- Messages may be traced as they progress through the opportunistic network, allowing an attacker to infer communication patterns [23]. The attacker would not necessarily need to be able to read the plaintext message body to perform this attack: eavesdropping on multiple encounters, and being able to detect that the same message was transmitted during each encounter (e.g., with transmission of the same encrypted text, or through matching headers), would be sufficient.
- Social network information is commonly used in opportunistic networks to inform routing decisions [24, 25, 26]. This social network information may be leaked via the routing scheme [27]. Learning the messages' header information, rather than the potentially end-to-end encrypted message bodies, is the target of this privacy threat.
- Locations of the participants may be inferred from the messages which their mobile devices carry — whether in absolute terms (“Alice is at the supermarket”), or relative terms (“Alice and Bob were in the same location this afternoon”) [27].

Some of these privacy concerns may be mitigated by technical measures, such as encryption. Even though a public key infrastructure may not be well-suited to opportunistic communications [28], decentralised mechanisms such as identity-based cryptography [22] — where the identity of each node acts as a key — may allow cryptographic solutions to some of the privacy risks. For example, if all pairs of nodes can communicate securely, then reading the contents of messages, or tracing message progression through the network by eavesdropping multiple encounters, becomes much more challenging.

But even identity-based cryptography requires a global trusted third party to vouch for new nodes entering the network (by generating the necessary private keys). This may not be feasible for certain types of opportunistic network such as those in rural or challenged environments, as access to the wider Internet may not be possible before encountering a new node. In such scenarios, it may be possible to employ simpler decentralised techniques, such as obfuscating information at the routing protocol or application layer [27], or utilising trusted social contacts [28].

3.3. *Location privacy*

In addition to the privacy threats caused by opportunistic networking, the use of location-based services over such a network introduces additional privacy con-

cerns. Location privacy has been studied in various contexts, such as sensor networks [29]; pervasive computing applications [30, 31]; and indeed opportunistic networks [32]. The focus of this research, however, has been on investigating the trade-offs involved in the protocols which work to preserve privacy, rather than considering how users’ own privacy-preserving behaviours may affect network performance, as we study in this paper.

Many user studies have been conducted to collect real-world data about users’ location privacy preferences [13, 33, 34]. More recently, the website “Please Rob Me”⁴ gathered publicly-available information from the social network sites Twitter and FourSquare in order to infer whether or not a person was at their home address [35] — further raising awareness of these potential privacy threats.

In the context of publishing sensed location information, various obfuscation techniques have been proposed [36, 37, 38]. User studies have also been performed to determine how users respond to various types of obfuscation [39]. But the focus has been on anonymity, rather than the performance of a distributed system.

4. User study methodology

To simulate mobile network performance in environments where users are subject to the potential privacy threats outlined in Section 3, we require a model of users’ privacy behaviour. There are two ways to obtain such a model. First, similarly to the commonly-used random waypoint mobility model [40], we could use a synthetic model generated according to mathematical properties. Such synthetic models are useful for simulation since they may be easy and tractable to generate, but may have limitations or rely on unrealistic assumptions, as outlined in Section 2. A second option, therefore, is to build a model based on measurements of actual user behaviour. To develop such a model, we measured privacy behaviour by performing a user study which investigated the location-sharing privacy preferences of 80 users of the popular Facebook SNS.

The overall purpose of our user study was to determine how widely participants would accept their current locations being broadcast to their friends online. Would a participant be happy to share some locations to the whole world, while others to selected friends only or to nobody at all? Are some participants more inclined to share their locations than other participants? Can we quantify location-sharing behaviours? Our second goal was to investigate the differences between a real and simulated network application, and so participants were divided into two groups,

⁴<http://pleaseroame.com/>

one of whom actually shared their locations with their Facebook friends, and one of whom did not.

Participants in the experiment carried a location-sensing mobile phone for one week of their day-to-day lives. Due to resource constraints — we had 20 mobile phones available, but 80 participants — the user study took place over four one-week runs, each with 20 participants. Two of the runs were conducted in St Andrews, a small town in the United Kingdom, and to ensure against bias from locale size, the other two runs were in London, the largest city in the UK. Participants were drawn from the undergraduate student population in both towns. We deliberately selected students who were not studying Computer Science (so that they would not be known by us), and who self-reported themselves as daily Facebook users. Further experimental details can be found in [41, 42].

Each participant was prompted up to 20 times per day to choose how widely their current location could be published on Facebook — to *everyone*, to some or all of their Facebook social contacts (“*friends*”), or to *nobody* at all. Optionally, they could ignore the prompts. A total of 4,334 prompts were sent, which resulted in 2,064 replies (a 47.6% response rate).

During each of the runs, half of the participants (10/20) were assigned at random to the *real* group, and half to the *simulation* group. Participants in the real group had their locations published to Facebook, visible to their social contacts according to their chosen preferences. Participants in the simulation group were able to see on Facebook the information which would have been published, but this was not disclosed to any of their social contacts.

We informed participants to which group they were assigned at the beginning of each run. Although we considered running a blind experiment, where participants would not be told to which group they were assigned, we decided to make the group explicit to participants. This is in order to allow comparison of the simulation group participants’ behaviour with that of participants in prior simulation-only studies, such as [13], where *all* participants used a simulated system, and were aware of this fact throughout.

By analogy to existing location-sharing applications, and the publicity surrounding information leakage (see Section 3.3), we believe that privacy choices for location-sharing behaviour when broadcasting locations via Facebook will not be dissimilar to those for an opportunistic network participant. As outlined in Section 3.2, one of the privacy risks associated with participation in an opportunistic network is the loss of location privacy. Should privacy threats due to information leakage in a real opportunistic network deployment receive similar publicity to similar threats in extant systems, then we believe that these location-sharing choices will converge to those for location broadcast, as we have measured in this user study. We can therefore use the results of this study to develop our privacy

model for opportunistic network behaviour.

5. A synthetic privacy model

As discussed in the previous section, we are interested in simulating privacy behaviour in mobile opportunistic networks. To first do this, we consider a *synthetic* privacy model. Such synthetic models are widely used in mobile networking research, such as the random waypoint mobility model. In this section, we extend the *epidemic routing* protocol [43] for opportunistic networks with a simple synthetic privacy model.

In epidemic routing the network is flooded with messages, by nodes exchanging messages in the message buffers of other nodes that they encounter, during each and every such encounter. We introduce a synthetic privacy model to this by assuming that nodes do not always exchange messages in every encounter, but would rather exchange messages during an encounter only with some fixed probability; this probability is a parameter of the synthetic privacy model. For example, with the probability parameter of 40%, in a given encounter two nodes would exchange the contents of their message buffers with one another with probability 40%; else no messages would be exchanged. Similarly, if the probability parameter was set to 0%, this would imply that nodes cared so much about their privacy that they refused to share data with any other nodes.

To evaluate the performance of opportunistic networks employing this synthetic privacy model with varying probability parameter, we use trace-driven simulation.

5.1. Datasets

We use two real-world datasets:

1. The *LocShare* dataset [44]. As described in Section 4, we collected the locations of 80 participants during four one-week runs of 20 participants. By defining an encounter as occurring when two participants are within 10 metres — selected as this is the approximate average Bluetooth range — we obtain a trace of encounters between participants. We use the Facebook lists of friends, as provided by participants, to determine a social network for each participant.
2. The *Reality Mining* dataset [45]. In this well-known dataset collected at MIT, 97 university members carried mobile phones during their daily lives over the course of an academic year. The phones recorded the results of periodic Bluetooth scans. As in [2, 27], we define Bluetooth encounters between

participant devices as opportunities for message exchange in an opportunistic network, and use each mobile phone’s address book to determine a social network for each user.

5.2. Simulation parameters

In line with our previous work [2], we use the following set of parameters for our simulations:

- Confidence intervals plotted based on 100 runs per data point.
- 100 messages per run.
- Unicast messages, sent from a sender node to one of the sender’s social network neighbours (*friends*). Note that although messages are unicast (destined for one particular recipient), the messages may follow multiple paths through the network in order to reach that destination.
- A message time-to-live (TTL) of one day.
- One week per simulation.⁵
- Infinite buffers and infinitely-fast transmission.⁶

We consider probability parameters for the synthetic model from 0% to 100%, in steps of 20%.

5.3. Results

Figures 1–2 show the performance of the network, as measured by two commonly-used metrics [24]:

- *Delivery ratio*: proportion of delivered messages, out of the total number of unique messages generated.
- *Delivery delay*: time taken for a message to first reach its destination.

⁵For *LocShare*, there are four one-week parts to the dataset; we therefore simulate 25 runs with each of the four one-week parts to make up the 100 runs for each datapoint. For *Reality Mining*, we pick a random one-week interval for each of the 100 runs — but, following [2, 27], we select only one-week intervals where there are sufficient numbers of nodes present for non-trivial routing to be possible.

⁶Our goal is to investigate the performance impact of privacy, so we do not set arbitrary constraints on buffer size or transmission rate, as these may confound the results.

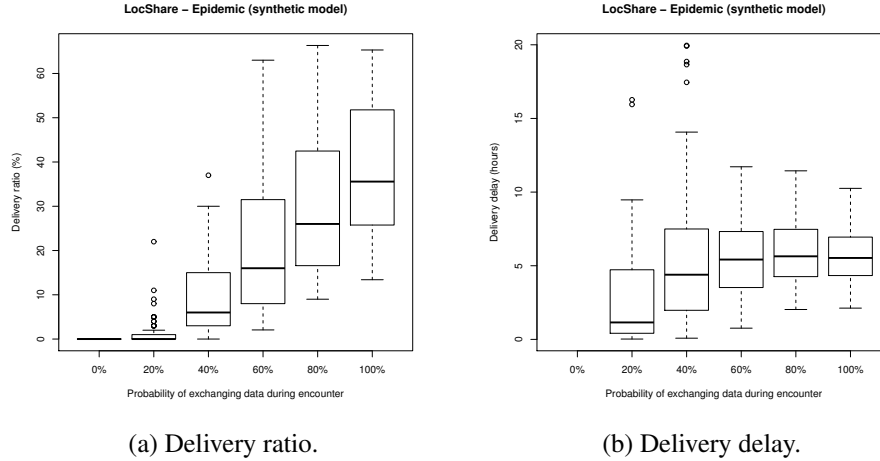


Figure 1: Delivery ratio and delay for the *LocShare* dataset under the synthetic privacy model. Delivery ratio varies widely based on the chosen probability parameter. Paradoxically, delivery delay is reduced when data are less likely to be exchanged during each encounter; successfully-delivered messages necessarily take shorter paths to their destination, since message exchange is unlikely during each encounter.

Figure 1a shows that routing performance, as measured by delivery ratio, for the *LocShare* dataset is significantly dependent on the probability parameter of the synthetic privacy model. As we might expect, when the probability parameter is 0%, messages are never exchanged during encounters, so the delivery delay falls to zero. Conversely, at 100%, performance is as for epidemic routing, since all possible messages are shared. Between these two extremes, there is a wide variation in the delivery ratio based on the parameter chosen, although it takes a value of 40% before any substantial message delivery is achieved.

Figure 1b shows that delivery delay is, paradoxically, reduced when messages are less likely to be exchanged during each encounter. By comparing with Figure 1a, we believe that this may be an artifact of the low delivery ratio: delivery delay is only considered for those messages which are successfully delivered. When the number of messages delivered is low, there is a survivor bias: the successfully-delivered messages necessarily take a shorter path to the destination, when message exchange is unlikely during each encounter.

A similar delivery ratio result holds for the *Reality Mining* dataset, as shown in Figure 2a. There is again a wide variation in delivery ratio, between the extremes of zero and epidemic routing's performance, based on the probability parameter.

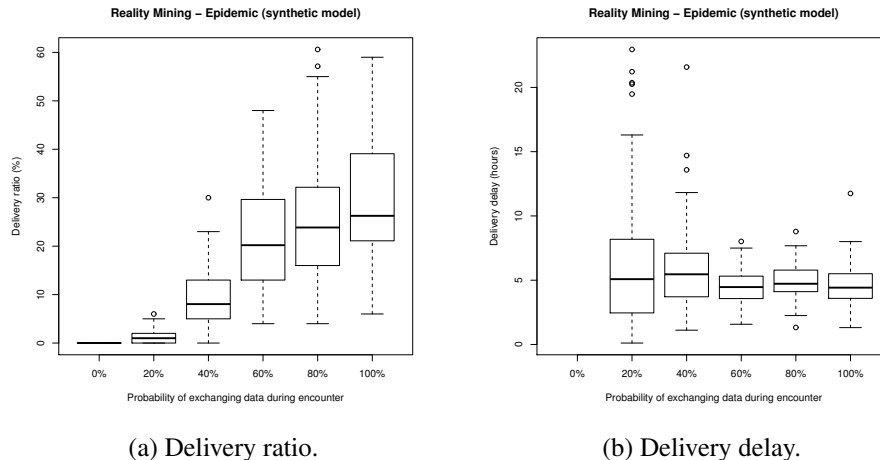


Figure 2: Delivery ratio and delay for the *Reality Mining* dataset under the synthetic model. Similarly to the *LocShare* dataset (Figure 1), there is a wide difference in delivery ratio based on the chosen probability parameter. Delivery delay does not significantly change. This may be an artifact of the sparse encounter trace: any successfully-delivered message will necessarily take approximately the same, small, number of hops.

The delivery delay, however, appears from Figure 2b to be independent of the probability parameter. A possible explanation is the sparseness of the *Reality Mining* dataset. Any successfully-delivered message would necessarily take approximately the same, small, number of hops — and so all messages would be delivered with similar delays. This may thus be a survivor bias, similar to Figure 1a, but as exhibited by a sparse dataset.

Since there is wide variation in the performance of the network based on the probability parameter for the synthetic model, what would be the performance of a real network? A limitation of the synthetic model is that there is no obvious criterion for choosing the probability parameter to use; delivery performance is not an intrinsic property of the dataset, but dependent on users’ propensity to share. In other words, performance will depend on how we model users’ privacy preferences. We therefore now consider a more realistic privacy model derived from our real-world user study.

6. An empirical privacy model

To obtain a privacy model that is more realistic than the simple synthetic model explored in Section 5, we use privacy preferences determined from real users dur-

Category	Proportion in category	Location sharing choice		
		Nobody	Friends	Everyone
Open	18% (7/40)	5.2%	7.7%	87.1%
Social	45% (18/40)	9.7%	80.5%	9.7%
Closed	28% (11/40)	72.1%	20.3%	7.6%
Variable	10% (4/40)	33.3%	29.6%	37.2%

Table 1: Real group ($n = 40$). Rows may not add up to 100% due to rounding.

Category	Proportion in category	Location sharing choice		
		Nobody	Friends	Everyone
Open	20% (8/40)	10.0%	3.9%	86.1%
Social	53% (21/40)	5.8%	75.4%	18.8%
Closed	18% (7/40)	67.0%	20.5%	12.6%
Variable	10% (4/40)	32.2%	36.4%	31.4%

Table 2: Simulation group ($n = 40$). Rows may not add up to 100% due to rounding.

ing the user study discussed in Section 4 to create an empirical privacy model. To study the differences between simulated and real applications, we further consider the participants in the real and simulation groups separately.

As is common in other privacy models [46], we segment the participants of each group into categories according to their privacy behaviour, i.e., their responses to the prompted questions (see Tables 1–2). We define four categories:

- *Open*: Participants usually shared their location publicly with everyone, in over 50% of responses.
- *Social*: Participants usually shared their location with some or all of their Facebook friends, in over 50% of responses.
- *Closed*: Participants usually did not share their location to anybody at all, in over 50% of responses.
- *Variable*: Participants did not have consistent location-sharing behaviour. They would sometimes share with nobody, with friends, and with everyone.

Note that, to a certain extent, all of our users had “variable” behaviour in that they did not act consistently at every location (we explore the implications of this further in [47]). Thus for the *Open*, *Social* and *Closed* groups, we consider a

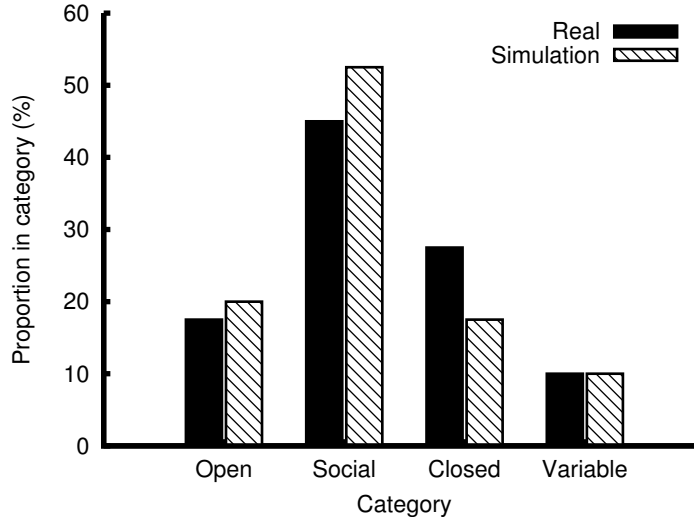


Figure 3: Proportion of participants in each category of the privacy model.

participant to be a member of this group if their behaviour is consistent with this group in over 50% of their sharing activity, i.e., their responses to questions.

For each of the four categories, we calculate the mean of the users’ location sharing choice proportions (*nobody*, *friends* or *everyone*) by user, in order to obtain Tables 1–2. Using these statistics, we can construct *privacy models* for users’ location sharing preferences, for each of the two groups of users. These privacy models are dataset-independent, and so may be applied to a variety of datasets for opportunistic network routing simulations. While it may be possible to construct more sophisticated models [17], our purpose here is to examine the differences in models between real and simulated applications, and so we limit ourselves to simple statistical models for now.

Figure 3 shows a visualisation of the relative number of participants in each of the four categories, for the real and simulation groups. The distribution of participants across the categories is generally similar between the real and simulation groups. As one might expect, most of our social network users were indeed social, and were willing to share with some or all of their Facebook friends. Smaller proportions were either more open or closed than this, or acted in a variable fashion. Note, however, that participants in the real group seemed more privacy-concerned than those in the simulation group: more participants in the real group were assigned to the privacy-concerned *Closed* category, and fewer to the less-privacy-concerned *Open* and *Social* categories, than in the simulation group. This might also be expected; participants might have taken more care about sharing their in-

formation if they knew that information was actually being shared on Facebook, and so acted in a more privacy-conscious, and perhaps more realistic, fashion.

We can now apply these privacy models for the two groups to opportunistic network simulations, in order to examine to what extent these differences in privacy behaviour affect routing performance.

7. Simulation results

To study the impact of simulated and real social networking applications on opportunistic network routing, we perform trace-driven simulations with the empirical privacy models, using the same datasets and simulation parameters as in Section 5.

At the start of each simulation run, each node (i.e., simulated participant) is allocated to one of the categories (*open*, *social*, *closed*, *variable*) for the duration of the run. We perform simulation for three methods of category allocation:

- *Central nodes closed*: Similarly to [48], we make use of the finding that the altruism (i.e., willingness to participate in message forwarding) of high-degree nodes is most important for network performance [49], to highlight any performance impact from the privacy models in our simulations. We rank nodes from highest to lowest degree-centrality in the encounter graph, and assign nodes in order to the *closed*, *social*, *variable* and *open* categories — with the number of nodes in each category according to the proportional size of the category in the privacy model. Higher-degree nodes are therefore less likely to forward messages, since they are assigned to the less open categories — which is analogous to behaving less altruistically — and so any performance impact due to the privacy model is maximised.
- *Central nodes open*: As a baseline, we perform simulations with the opposite allocation strategy to *central nodes open*. The nodes are ranked from *lowest to highest* degree-centrality (the reverse ordering to *Central nodes closed*), and then allocated to categories as before. The lowest-degree nodes are then less likely to forward messages, while the highest-degree nodes are more likely to do so.
- *Random category allocation*: As a further baseline, nodes are assigned randomly to categories, with allocation probability proportional to the size of each category. Centrality is not considered in the allocation process.

Algorithm 1 shows pseudocode for these three category allocation methods.

Algorithm 1 Nodes' category allocation

```
1: if allocation_scheme == 'random' then
2:   for all node in nodes do
3:     node_category  $\leftarrow$  random category (weighted by category size)
4: else
5:   ordered_nodes_stack  $\leftarrow$  []
6:   if allocation_scheme == 'central nodes closed' then
7:     ordered_nodes_stack  $\leftarrow$  [nodes ordered by descending centrality]
8:   else if allocation_scheme == 'central nodes open' then
9:     ordered_nodes_stack  $\leftarrow$  [nodes ordered by ascending centrality]
10:  for all category in categories do
11:    wanted_category_size[category]  $\leftarrow$ 
12:    num_nodes * proportional_size[category]
13:    while node  $\leftarrow$  pop ordered_nodes do
14:      for all category in [closed, social, variable, open] do
15:        if num_in_category[category] < wanted_category_size[category]
16:        && node_category is unassigned then
17:          node_category  $\leftarrow$  category
```

Following [2], we define two modes of privacy behaviour to apply the privacy models for the real and simulation groups to our opportunistic network simulations. While it is possible to think of many more behaviours, we believe that two modes are sufficient for investigating the impact of privacy. Previous work has also demonstrated that a constrained number of privacy choices is a usable compromise for privacy policies for ubiquitous computing environments [50]. Our chosen modes are:

- *Friendly (F)*: Nodes are modelled as being willing to share with their social network friends. If the overall privacy choice is *everyone*, then the nodes behave as in the default case; if *nobody*, then messages are not exchanged; if *friends*, then as the default case only if the two nodes involved in this encounter are friends (otherwise messages are not exchanged).
- *PubPriv (PP)*: Nodes are modelled as either being fully public (no privacy concerns), or fully private (any privacy concerns result in disregarding the encounter) — with nothing in-between. If the overall privacy behaviour during an encounter is *everyone*, then messages are exchanged as in the default case. Otherwise (i.e., if the overall privacy behaviour is *friends* or *nobody*), messages are not exchanged.

During each encounter between a pair of nodes, each of the two nodes randomly picks a privacy behaviour of $\{nobody, friends, everyone\}$, weighted according to the location-sharing proportions associated with that node’s category. Messages are then exchanged depending on the chosen privacy behaviours for that encounter. The overriding choice is the more restrictive of the two nodes’ privacy behaviours. For example, if one node picks *nobody* and the other picks *everyone*, then the overall choice is the more restrictive *nobody*.

Algorithm 2 shows pseudocode for the logic behind choosing whether messages are exchanged between nodes during each encounter, for each of the privacy modes.

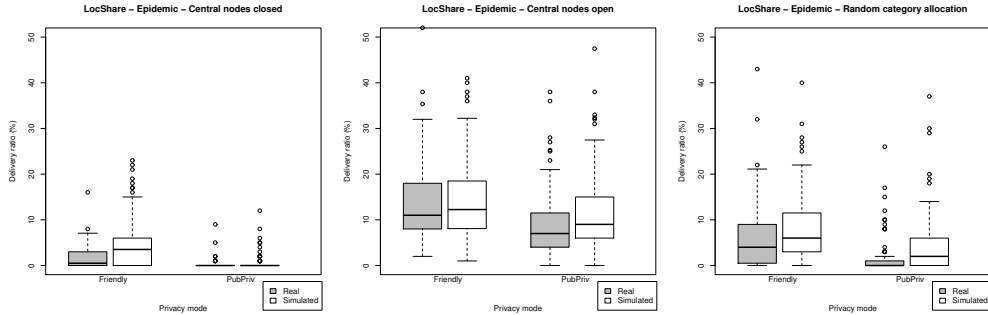
Algorithm 2 Decision: exchange messages during encounter

procedure Encounter($node_A, node_B$)

- 1: {The weighting of the *privacy_behaviour* random choice is specified by the privacy model. Random weighting according to each node’s previously allocated category.}
 - 2: $privacy_behaviour[node_A] \leftarrow$ random choice of $\{nobody, friends, everyone\}$ weighted by weightings for $node_category[node_A]$
 - 3: $privacy_behaviour[node_B] \leftarrow$ random choice of $\{nobody, friends, everyone\}$ weighted by weightings for $node_category[node_B]$
 - 4: $overall_privacy_behaviour \leftarrow$ more restrictive privacy choice of $node_A$ or $node_B$
 - 5: **if** $mode == Friendly$ **then**
 - 6: **if** $(overall_privacy_behaviour == everyone)$ ||
 $(overall_privacy_behaviour == friends \ \&\& \ are_friends(node_A, node_B))$
 then
 - 7: messages are exchanged
 - 8: **else if** $mode == PubPriv$ **then**
 - 9: **if** $overall_privacy_behaviour == everyone$ **then**
 - 10: messages are exchanged
-

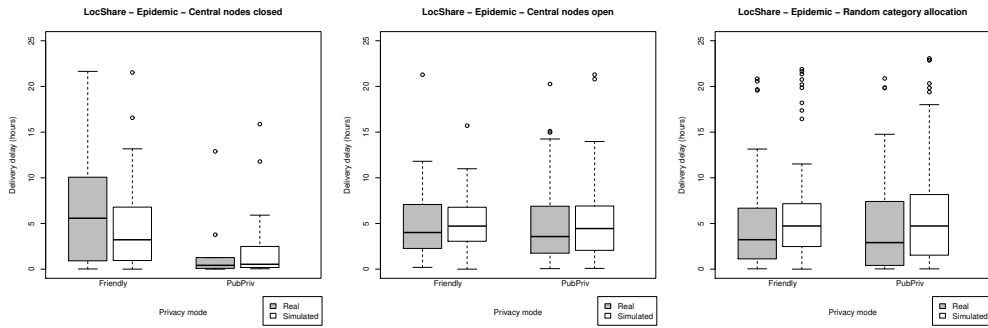
Figure 4a shows that, when magnifying differences between privacy models with the *Central nodes closed* allocation scheme, the Friendly mode delivery ratio for the real group model is significantly lower than for the simulation group’s model (0.5% vs 3.5%).⁷ For the stricter *PubPriv* mode, the median delivery ratio

⁷Note that these low delivery ratios, in absolute terms, are typical of opportunistic network routing simulations: the datasets used, while collected from real-world users, are from experimental settings that are not as highly-connected or dense as we would expect in a real large-scale deployment.



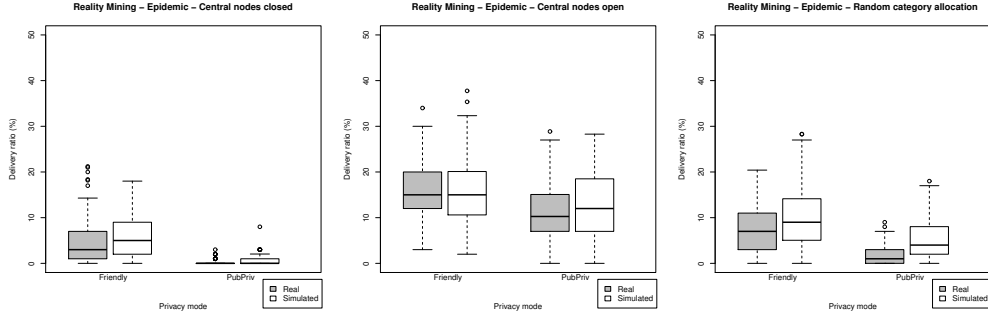
(a) *Central nodes closed.* (b) *Central nodes open.* (c) *Random.*

Figure 4: *LocShare* dataset: Delivery ratios for the real group and the simulated group, for two privacy models under three allocation methods. When magnifying differences (the *Central nodes closed* allocation scheme), under the Friendly mode, the delivery ratio for the real group model is significantly lower than for the simulation group model (0.5% vs 3.5%). Under the stricter PubPriv model, the delivery ratio for both groups is zero. As expected, when central nodes act increasingly openly in the baseline simulations, absolute delivery ratios increase. For these baseline simulations, delivery ratios are generally similar across real and simulation groups’ models — except for the PubPriv mode under the *Random* allocation, where a similar trend is visible (0.0% vs 2.0%).



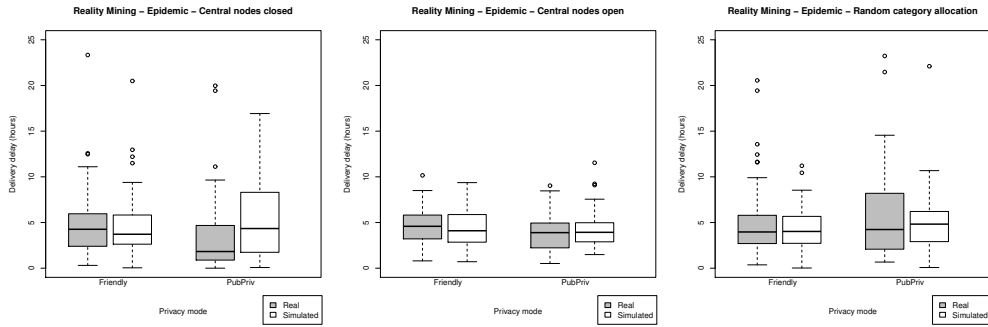
(a) *Central nodes closed.* (b) *Central nodes open.* (c) *Random.*

Figure 5: *LocShare* dataset: Delivery delay comparison between the privacy models of the real and simulated groups. The trend is similar across all three category allocation methods: there is no significant difference in delivery delay between the real and simulation group privacy models, with wide overlap between the bars. Since delivery delay is only measured for the few messages that arrive, paradoxically the delay may be lower for the PubPriv mode than the Friendly mode for both groups’ privacy models, but is similar between models.



(a) *Central nodes closed.* (b) *Central nodes open.* (c) *Random.*

Figure 6: *Reality Mining* dataset: Delivery ratio comparison between the privacy models of the real and simulated groups. Performance is generally similar between the real and simulation groups' privacy models, with large overlap between the boxes. A difference is pronounced only for the PubPriv dataset in the *Random* allocation scheme, as seen with the *LocShare* dataset.



(a) *Central nodes closed.* (b) *Central nodes open.* (c) *Random.*

Figure 7: *Reality Mining* dataset: Delivery delay comparison between the privacy models of the real and simulated groups. As for the *LocShare* dataset, the delivery delay is not significantly changed between the real and simulation groups, for either mode across any of the three allocation methods — with the exception of the paradoxically low PubPriv delivery delay for the real group under the *Central nodes closed* allocation scheme, due to the low delivery ratios.

under both models is zero. As would be expected, when central nodes act increasingly openly in the baseline simulations (Figures 4b–4c), the absolute delivery ratios increase. For these baseline simulations, the delivery ratios are generally similar across real and simulation groups’ models — except for the PubPriv mode under the Random allocation, where a similar trend is visible (0.0% vs 2.0%).

There is a wide variation in delivery delay, as shown in Figures 5a–5c. But performance is not significantly different between the real and simulation group privacy models, with wide overlap between the bars for each privacy mode and allocation scheme. Only when magnifying differences, in Figure 5a, is performance is again poorer under this metric for the simulations utilising the real group’s privacy model: under the *Friendly* privacy mode, median delivery delay rises from 3 hours for the simulation group to 6 hours for the real group. For the *PubPriv* privacy mode the delivery ratios are so low that, as for Figure 3, the corresponding delivery delays are paradoxically low.

For the *Reality Mining* dataset, Figure 6a shows that the differences in delivery ratio between corresponding simulations utilising the real group and simulation groups’ privacy models are relatively smaller: the boxes overlap. As for the *LocShare* dataset, however, the median delivery ratio under the *Friendly* mode is lower for the real group as compared to the simulation group (3% vs 5%), and for the *PubPriv* mode is zero in both cases. The baseline simulations (Figures 6b–6c) display the same trend, with similar performance for the real and simulation groups’ privacy modes. Only for the PubPriv dataset in the Random allocation scheme is there a significant difference (1.0% vs 4.0%), as for the *LocShare* dataset.

Correspondingly, delivery delay is not significantly changed between the real and simulation group models: Figures 7a–7c show that the median delivery delay is constant at approximately four hours for each of the real and simulation groups’ privacy models, for each privacy mode and category allocation scheme. The one seemingly significant difference, for the *PubPriv* mode in Figure 7a, is, as for the *LocShare* dataset, an artifact of the low delivery ratios.

8. Conclusions and future work

Simulation is a useful method for developing future network protocols where the construction and modification of such networks may prove impractical. But future networks may also be used by future applications, which may similarly not exist and therefore also need simulation. If assumptions about user behaviour in these applications are unrealistic, then this may have an impact on the credibility of any simulation results.

In this paper we have studied the interplay between real and simulated applications in user studies and their effect on simulations at the network layer, in partic-

ular in a mobile opportunistic network. First, we find that assumptions about user behaviour do indeed have a huge impact on network performance. If we consider privacy behaviour, performance (as measured by delivery ratio) in an opportunistic network can fall to zero, implying that the network is useless. Second, different assumptions about behaviour can also have an impact. When we compare a synthetic privacy model to an empirically-determined privacy model derived from real user behaviour in a user study, we find dramatically different levels of network performance. Finally, we find that models derived from real and simulated applications do not seem to differ greatly in their impact on network performance. When performing simulations without magnifying any differences between these privacy models by choosing well-connected nodes with high degrees to act more privacy-concerned, we did not observe any significant difference in network performance across the real and simulation group models. We did, however, observe differences based on the privacy model, and the allocation scheme for which nodes should be considered privacy-concerned. We believe that simulated applications may thus be used as a reasonable predictor of how users' privacy behaviour may affect networks.

Our results suggest some simple guidelines for researchers who wish to conduct realistic network simulations:

- First, use models based on real user behaviour to drive application-layer behaviour in network-layer simulations.
- Second, if real applications do not exist or are impractical, use simulated applications to obtain information about user behaviour.

The empirical privacy models that we present have been developed to be dataset-independent, and we have applied them to two real-world traces. That said, we have not tested whether our results generalise and so our guidelines should be taken with caution. In fact, it may not be possible to exhaustively test whether our results generalise, as this would involve a huge number of user studies with a variety of applications in a variety of real-world scenarios.

In this paper we present results for one routing protocol, *epidemic*, with added privacy models. We have also performed similar simulations with another routing protocol, *simple social network routing* [2], and observed the same trends, but due to space constraints, we have not included the plots in this paper.

We have studied a particular application (a mobile social location-sharing application), and a particular type of network (a mobile ad hoc opportunistic network). Future work will be needed to conduct user studies in other types of real and simulated applications, and models compared and applied to different network types.

In addition to further applications and types of network, our existing dataset requires further investigation. We plan to investigate more sophisticated privacy models. For instance, is there a correlation between privacy preferences and the location of an encounter? We have also assumed that the privacy behaviour of heavy Facebook users corresponds to that of (potentially-pseudonymous) opportunistic network users. Future work needs to be conducted to test whether this assumption holds. At the same time, more complex scenarios might also change behaviour; for example, other privacy modes beyond the two that we identified. Moreover, more complex privacy-preserving protocols, which may alleviate users' location privacy concerns (and hence reduce the impact on performance from user behaviour) by preserving privacy, but potentially at a performance cost from the protocol itself, might also change user's privacy behaviour.

Future work may also consider performance of the network, on adding privacy modes, under metrics other than delivery ratio and delay. Since delivery ratios were so low, we found that — paradoxically — reduced performance under the metric of delivery ratio may lead to seemingly-improved performance under the metric of delivery delay.

A potential limitation of our privacy models is that we asked users for their preferences to disclose their *exact* location — to the accuracy of a GPS sensor. Should location information be leaked via participation in opportunistic networks, then perhaps only coarser locations may be discoverable by other users. Since coarse locations introduce a degree of obfuscation, thus implicitly providing users with increased privacy [39], we would like to investigate how user preferences may vary depending on the granularity of their shared location, within the context of an opportunistic networking application.

On the other hand, a strength of our privacy models is that they are dataset-independent, and so may be applied to other datasets, or scenarios beyond opportunistic network routing. So it would be possible for researchers to use these models as input to other performance evaluations, to build upon our evaluation.

If users' location privacy concerns can have such an impact on opportunistic network performance, might other privacy concerns (for example, about some malicious node gathering and broadcasting information, or the possibility of the revelation of users' friends lists [27]) also have a performance impact? To find out, we need to obtain reliable data on other types of privacy concerns in these networks. More data on encounters would also be useful: existing datasets used to evaluate opportunistic network protocols are relatively small-scale compared to the population of a town or city; perhaps evaluation against a larger-scale (and hence perhaps more-highly-connected) dataset would yield new insights. We are searching for such datasets. We are sharing the dataset used here through the CRAWDAD wireless data archive [44] and encourage other researchers to do the same so that

we can further evaluate our protocols.

Acknowledgements

This work was supported by the Engineering and Physical Sciences Research Council [grant number EP/G002606/1].

References

- [1] I. Parris, F. Ben Abdesslem, T. Henderson, Facebook or Fakebook?: The effect of simulation on location privacy user studies, in: Proceedings of the Privacy and Usability Methods Pow-Wow (PUMP), British Computer Society, 2010.
URL <http://scone.cs.st-andrews.ac.uk/pump2010/papers/parris.pdf>
- [2] I. Parris, T. Henderson, The impact of location privacy on opportunistic networks, in: Proceedings of the Fifth IEEE WoWMoM Workshop on Autonomic and Opportunistic Communications (AOC), IEEE, 2011. doi:10.1109/WoWMoM.2011.5986149.
- [3] T. R. Andel, A. Yasinac, On the credibility of MANET simulations, IEEE Computer 39 (7) (2006) 48–54. doi:10.1109/MC.2006.242.
- [4] C. Newport, D. Kotz, Y. Yuan, R. S. Gray, J. Liu, C. Elliott, Experimental evaluation of wireless simulation assumptions, SIMULATION 83 (9) (2007) 643–661. doi:10.1177/0037549707085632.
- [5] K. Tan, D. Wu, A. Jack Chan, P. Mohapatra, Comparing simulation tools and experimental testbeds for wireless mesh networks, in: IEEE International Symposium on A World of Wireless, Mobile and Multimedia Networks (WoWMoM), IEEE, 2010. doi:10.1109/WOWMOM.2010.5534917.
- [6] A. Torres, C. T. Calafate, J.-C. Cano, P. Manzoni, Deploying a real IEEE 802.11e testbed to validate simulation results, in: Proceedings of the IEEE 34th Conference on Local Computer Networks (LCN 2009), 2009, pp. 109–115. doi:10.1109/LCN.2009.5355191.
- [7] S. Kurkowski, T. Camp, M. Colagrosso, MANET simulation studies: the incredibles, ACM SIGMOBILE Mobile Computing and Communications Review 9 (4) (2005) 50–61. doi:10.1145/1096166.1096174.

- [8] K. Pawlikowski, H. D. J. Jeong, J. S. R. Lee, On credibility of simulation studies of telecommunication networks, *IEEE Communications Magazine* 40 (1) (2002) 132–139. doi:10.1109/35.978060.
- [9] S. Floyd, V. Paxson, Difficulties in simulating the Internet, *IEEE/ACM Transactions on Networking* 9 (4) (2001) 392–403. doi:10.1109/90.944338.
- [10] R. Chertov, S. Fahmy, Forwarding devices: From measurements to simulations, *ACM Transactions on Modeling and Computer Simulation* 21 (2). doi:10.1145/1899396.1899400.
- [11] Y. Kim, K. Taylor, C. Dunbar, B. Walker, P. Mundur, Reality vs emulation: running real mobility traces on a mobile wireless testbed, in: *Proceedings of the 3rd ACM international workshop on Hot Topics in Planet-Scale Measurement*, ACM, New York, NY, USA, 2011, pp. 23–28. doi:10.1145/2000172.2000180.
- [12] J. Y. Tsai, P. Kelley, P. Drielsma, L. F. Cranor, J. Hong, N. Sadeh, Who’s viewed you?: The impact of feedback in a mobile location-sharing application, in: *CHI ’09: Proceedings of the 27th international conference on Human factors in computing systems*, ACM, New York, NY, USA, 2009, pp. 2003–2012. doi:10.1145/1518701.1519005.
- [13] S. Consolvo, I. E. Smith, T. Matthews, A. Lamarca, J. Tabert, P. Powledge, Location disclosure to social relations: why, when, & what people want to share, in: *CHI ’05: Proceedings of the 23rd international conference on Human factors in computing systems*, ACM Press, New York, NY, USA, 2005, pp. 81–90. doi:10.1145/1054972.1054985.
- [14] S. Consolvo, M. Walker, Using the experience sampling method to evaluate ubicomp applications, *IEEE Pervasive Computing* 2 (2) (2003) 24–31. doi:10.1109/MPRV.2003.1203750.
- [15] F. Stutzman, J. K. Duffield, Friends only: examining a privacy-enhancing behavior in facebook, in: *CHI ’10: Proceedings of the 28th international conference on Human factors in computing systems*, ACM, New York, NY, USA, 2010, pp. 1553–1562. doi:10.1145/1753326.1753559.
- [16] A. L. Young, A. Quan-Haase, Information revelation and internet privacy concerns on social network sites: a case study of Facebook, in: *C&T ’09: Proceedings of the fourth international conference on Communities and technologies*, ACM Press, New York, NY, USA, 2009, pp. 265–274. doi:10.1145/1556460.1556499.

- [17] E. Toch, J. Cranshaw, P. H. Drielsma, J. Y. Tsai, P. G. Kelley, J. Springfield, L. Cranor, J. Hong, N. Sadeh, Empirical models of privacy in location sharing, in: *Proceedings of the 12th ACM international conference on Ubiquitous computing, Ubicomp '10*, ACM, New York, NY, USA, 2010, pp. 129–138. doi:10.1145/1864349.1864364.
- [18] G. K. Orman, V. Labatut, The effect of network realism on community detection algorithms, in: *2010 International Conference on Advances in Social Networks Analysis and Mining*, IEEE, 2010, pp. 301–305. doi:10.1109/ASONAM.2010.70.
- [19] A. K. Pietiläinen, E. Oliver, J. Lebrun, G. Varghese, C. Diot, MobiClique: middleware for mobile social networking, in: *WOSN '09: Proceedings of the 2nd ACM Workshop on Online Social Networks*, ACM, New York, NY, USA, 2009, pp. 49–54. doi:10.1145/1592665.1592678.
- [20] L. Pelusi, A. Passarella, M. Conti, Opportunistic networking: data forwarding in disconnected mobile ad hoc networks, *IEEE Communications Magazine* 44 (11) (2006) 134–141. doi:10.1109/MCOM.2006.248176.
- [21] J. Crowcroft, E. Yoneki, P. Hui, T. Henderson, Promoting tolerance for delay tolerant network research, *ACM SIGCOMM Computer Communication Review* 38 (5) (2008) 63–68. doi:10.1145/1452335.1452345.
- [22] A. Shikfa, M. Önen, R. Molva, Privacy and confidentiality in context-based and epidemic forwarding, *Computer Communications* 33 (13) (2010) 1493–1504. doi:10.1016/j.comcom.2010.04.035.
- [23] Z. Le, G. Vakde, M. Wright, PEON: privacy-enhanced opportunistic networks with applications in assistive environments, in: *PETRA '09: Proceedings of the 2nd International Conference on Pervasive Technologies Related to Assistive Environments*, ACM Press, New York, NY, USA, 2009, pp. 1–8. doi:10.1145/1579114.1579190.
- [24] P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: social-based forwarding in delay tolerant networks, in: *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing (MobiHoc)*, ACM, New York, NY, USA, 2008, pp. 241–250. doi:10.1145/1374618.1374652.
- [25] E. M. Daly, M. Haahr, Social network analysis for information flow in disconnected delay-tolerant MANETs, *IEEE Transactions on Mobile Computing* 8 (5) (2009) 606–621. doi:10.1109/TMC.2008.161.

- [26] C. Boldrini, M. Conti, A. Passarella, Exploiting users' social relations to forward data in opportunistic networks: The HiBOp solution, *Pervasive and Mobile Computing* 4 (5) (2008) 633–657. doi:10.1016/j.pmcj.2008.04.003.
- [27] I. Parris, T. Henderson, Privacy-enhanced social-network routing, *Computer Communications* 35 (1) (2012) 62–74. doi:10.1016/j.comcom.2010.11.003.
- [28] K. El Defrawy, J. Solis, G. Tsudik, Leveraging social contacts for message confidentiality in delay tolerant networks, in: 2009 33rd Annual IEEE International Computer Software and Applications Conference, IEEE, 2009, pp. 271–279. doi:10.1109/COMPSAC.2009.43.
- [29] P. Kamat, Y. Zhang, W. Trappe, C. Ozturk, Enhancing Source-Location privacy in sensor network routing, in: Proceedings of the 25th IEEE International Conference on Distributed Computing Systems, Columbus, OH, USA, 2005, pp. 599–608. doi:10.1109/ICDCS.2005.31.
- [30] A. R. Beresford, F. Stajano, Location privacy in pervasive computing, *IEEE Pervasive Computing* 2 (1) (2003) 46–55. doi:10.1109/MPRV.2003.1186725.
- [31] M. Xue, P. Kalnis, H. K. Pung, Location diversity: Enhanced privacy protection in location based services, in: Proceedings of the 4th International Symposium on Location and Context Awareness (LOCA), Vol. 5561 of Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, 2009, pp. 70–87. doi:10.1007/978-3-642-01721-6_5.
- [32] X. Lu, P. Hui, D. Towsley, J. Pu, Z. Xiong, Anti-localization anonymous routing for delay tolerant network, *Computer Networks* 54 (11) (2010) 1899–1910. doi:10.1016/j.comnet.2010.03.002.
- [33] M. Benisch, P. Kelley, N. Sadeh, L. Cranor, Capturing location-privacy preferences: quantifying accuracy and user-burden tradeoffs, *Personal and Ubiquitous Computing* (2010) 1–16doi:10.1007/s00779-010-0346-0.
- [34] D. Anthony, T. Henderson, D. Kotz, Privacy in Location-Aware computing environments, *IEEE Pervasive Computing* 6 (4) (2007) 64–72. doi:10.1109/MPRV.2007.83.
- [35] S. Gambs, M. O. Killijian, M. N. nez del Prado Cortez, Show me how you move and I will tell you who you are, in: Proceedings of

the 3rd ACM SIGSPATIAL International Workshop on Security and Privacy in GIS and LBS, ACM, New York, NY, USA, 2010, pp. 34–41. doi:10.1145/1868470.1868479.

- [36] C. Ardagna, M. Cremonini, E. Damiani, S. De Capitani di Vimercati, P. Samarati, Location privacy protection through Obfuscation-Based techniques, *Data and Applications Security XXI (2007)* 47–60doi:10.1007/978-3-540-73538-0_4.
- [37] J. Krumm, Realistic Driving Trips For Location Privacy, in: *Proceedings of the Seventh International Conference on Pervasive Computing*, Vol. 5538 of *Lecture Notes in Computer Science*, Springer Berlin / Heidelberg, Berlin, Heidelberg, 2009, pp. 25–41. doi:10.1007/978-3-642-01516-8_4.
- [38] G. Zhong, U. Hengartner, Toward a distributed k-anonymity protocol for location privacy, in: *WPES 2008: Proceedings of the 7th ACM workshop on Privacy in the electronic society*, ACM, New York, NY, USA, 2008, pp. 33–38. doi:10.1145/1456403.1456410.
- [39] A. J. B. Brush, J. Krumm, J. Scott, Exploring end user preferences for location obfuscation, location-based services, and the value of location, in: *Ubicomp '10: Proceedings of the 12th ACM international conference on Ubiquitous computing*, ACM, New York, NY, USA, 2010, pp. 95–104. doi:10.1145/1864349.1864381.
- [40] D. B. Johnson, D. A. Maltz, Dynamic source routing in ad hoc wireless networks, in: T. Imielinski, H. F. Korth (Eds.), *Mobile Computing*, Vol. 353 of *The Kluwer International Series in Engineering and Computer Science*, Springer US, Boston, MA, USA, 1996, Ch. 5, pp. 153–181. doi:10.1007/978-0-585-29603-6_5.
- [41] F. Ben Abdesslem, I. Parris, T. Henderson, Mobile experience sampling: Reaching the parts of Facebook other methods cannot reach, in: *Proceedings of the Privacy and Usability Methods Pow-Wow (PUMP)*, British Computer Society, 2010.
URL <http://scone.cs.st-andrews.ac.uk/pump2010/papers/benabdesslem.pdf>
- [42] F. Ben Abdesslem, I. Parris, T. Henderson, Reliable online social network data collection, in: A. Abraham, A. Ella Hassanien (Eds.), *Computational Social Networks: Mining and Visualization*, Vol. 3 of *Springer Computer Communications and Networks Series*, Springer-Verlag, London, UK, 2011, accepted for publication.

- [43] A. Vahdat, D. Becker, Epidemic Routing for Partially-Connected Ad Hoc Networks, Tech. Rep. CS-200006, Duke University (Apr. 2000).
URL <http://issg.cs.duke.edu/epidemic/epidemic.pdf>
- [44] F. Ben Abdesslem, T. Henderson, I. Parris, CRAWDAD data set st_andrews/locshare (v. 2011-10-12), Downloaded from http://crawdad.org/st_andrews/locshare/ (Oct. 2011).
- [45] N. Eagle, A. S. Pentland, D. Lazer, Inferring friendship network structure by using mobile phone data, *Proceedings of the National Academy of Sciences* 106 (36) (2009) 15274–15278. doi:10.1073/pnas.0900282106.
- [46] A. F. Westin, Social and political dimensions of privacy, *Journal of Social Issues* 59 (2) (2003) 431–453. doi:10.1111/1540-4560.00072.
- [47] F. Ben Abdesslem, T. Henderson, S. Brostoff, M. A. Sasse, Context-based personalised settings for mobile location sharing, in: *Proceedings of the ACM Recommender Systems Workshop on Personalization in Mobile Applications*, ACM, 2011.
- [48] G. Bigwood, T. Henderson, IRONMAN: Using social networks to add incentives and reputation to opportunistic networks, in: *Proceedings of the IEEE Third International Conference on Social Computing (SocialCom)*, IEEE, 2011, pp. 65–72. doi:10.1109/PASSAT/SocialCom.2011.60.
- [49] K. Xu, P. Hui, V. O. Li, J. Crowcroft, V. Latora, P. Lio, Impact of altruism on opportunistic communications, in: *Proceedings of the First International Conference on Ubiquitous and Future Networks (ICUFN)*, IEEE, 2009, pp. 153–158. doi:10.1109/ICUFN.2009.5174303.
- [50] A. Kapadia, T. Henderson, J. Fielding, D. Kotz, Virtual Walls: Protecting Digital Privacy in Pervasive Environments, in: *Proceedings of the 5th International Conference on Pervasive Computing*, no. 4480 in LNCS, Toronto, Canada, 2007, pp. 162–179. doi:10.1007/978-3-540-72037-9_10.