

Practical privacy-aware opportunistic networking

Iain Parris
School of Computer Science
University of St Andrews
St Andrews, Fife, KY16 9SX, UK
ip@cs.st-andrews.ac.uk

Tristan Henderson
School of Computer Science
University of St Andrews
St Andrews, Fife, KY16 9SX, UK
tristan@cs.st-andrews.ac.uk

Opportunistic networks have been the study of much research — in particular on making end-to-end routing efficient. Users’ privacy concerns, however, have not been the subject of much research. What privacy concerns might opportunistic network users have? Is it possible to build opportunistic networks that can mitigate users’ privacy concerns while maintaining routing performance?

Our work-to-date has tackled the problem of creating privacy-preserving routing protocols, with less emphasis on discovering users’ actual privacy concerns. We summarise our current results, and describe a future experiment that we have planned to better understand users’ privacy concerns.

opportunistic networking, privacy, social networks, routing protocols

1. INTRODUCTION

People commonly carry mobile devices — such as phones — during their daily lives. When in proximity, these devices may exchange data directly, without using any traditional infrastructure, via a protocol such as Bluetooth. If many such devices cooperate with one another, an *opportunistic network* may be formed (Pelusi et al. 2006). Data are exchanged between mobile devices opportunistically as they move into physical proximity, in a disconnected store-and-forward architecture.

There are a number of challenges in opportunistic networking research. One challenge is *routing*. Given episodic connectivity, based on people’s real-world movements, how might we efficiently route messages through the network? If we naively exchange messages during each and every encounter, flooding messages out along all possible paths, then the message will certainly find and follow any existing path — indeed, the shortest path — between sender and destination to be delivered as quickly as possible.¹ But this *epidemic routing* approach is costly: large numbers of redundant messages are typically sent, which may rapidly drain the batteries of the mobile devices. Therefore, various routing schemes have been proposed that utilise social network information to inform routing decisions. (Hui et al. 2008; Daly and Haahr 2009; Boldrini et al. 2008).

A second, related challenge is *privacy*. Through participating in an opportunistic network — and especially

if social network information is used to inform routing decisions — users may experience a variety of privacy threats (Parris and Henderson 2011b).

Our research focuses on the intersection of these challenges. What privacy concerns might opportunistic network users have? How might we measure these concerns for future application users, given the disconnect between what people say and what people do with respect to privacy? Is it possible to build opportunistic networks which can mitigate users’ privacy concerns while maintaining routing performance?

The proposed PhD thesis statement is: *it is possible to maintain opportunistic network performance after adding the privacy-preserving features that users desire.*

Our work-to-date has tackled the problem of creating privacy-preserving protocols. We have performed simulations with real users’ data — location traces, and location-privacy preferences — to quantify the impact of privacy preferences on performance.

Our focus is now on a problem up-the-stack: the user. We wish to perform an HCI experiment, utilising a new methodology, to measure the privacy concerns of users for an example future opportunistic network application. Ultimately, the goal is to improve the happiness of users, through securing their potentially-sensitive data — which we note ties in with the HCI 2011 *Health, wealth & happiness* theme. We hope to receive useful feedback and suggestions at the HCI 2011 Doctoral Consortium.

¹Under ideal conditions. If storage space is finite, for example, then nodes may run out of storage space and drop messages, and thus this may not be the case.

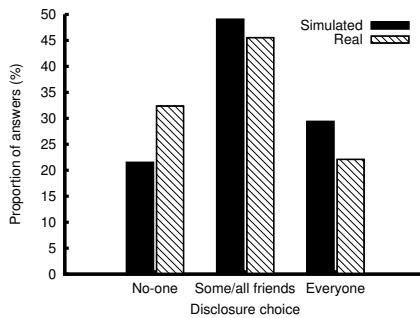


Figure 1: Location-sharing preferences of Facebook users. Those experiencing simulated publishing of their data share their data more openly than those experiencing real publishing: they less often disclose their location to no-one, and more often to everyone.

2. CURRENT RESULTS

So far, we have found three main results:

1. Users behave differently when participating in a real system, compared to a simulated system.
2. Users’ location-privacy preferences may significantly impact opportunistic-network performance.
3. Significant obfuscation of social-network information is possible, while still maintaining good social-network routing performance.

We describe each in turn.

2.1. User study: location-privacy preferences

We performed a user study, investigating the location-sharing privacy preferences of 80 users of the popular online social network Facebook.²

Participants carried a location-sensing mobile phone for one week of their daily lives. Due to resource constraints — we had 20 mobile phones available, but 80 participants — we performed the user study in four one-week runs, each with 20 participants. Two of the runs were conducted in a small UK town, St Andrews; the other two runs were conducted in a large UK city, London.

Utilising the experience sampling method (Consolvo and Walker 2003), participants were prompted up to 20 times each day to choose how widely they would be happy for their current location to be shared on Facebook — to everyone, to some or all of their Facebook social contacts, or to no-one.

At the start of each of the four runs, the 20 participants in the run were randomly divided into two groups. The *real group* experienced real publishing of their location information on Facebook; the *simulation group*

²<http://www.facebook.com/>

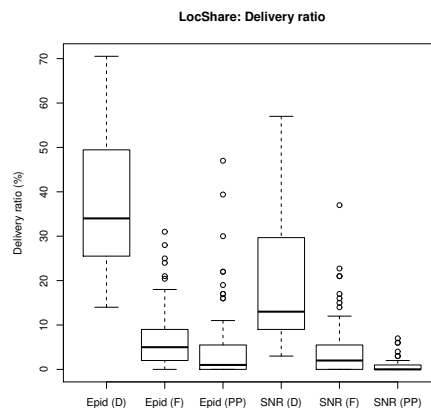


Figure 2: Delivery ratio (i.e., proportion of unique messages that are delivered) for epidemic and social-network routing, under three different privacy modes. The delivery ratio falls significantly when privacy concerns are taken into account (the Friendly (F) and PubPriv (PP) privacy modes).

experienced simulated publishing, where information was never disclosed to anybody, regardless of user preferences.

We investigated whether publishing information “for real” (the real group) resulted in a difference of behaviour compared to simulated publishing. Figure 1 shows our main result: *the simulation group shared their locations more openly than the real group* (Parris et al. 2010).³

2.2. Performance impact of users’ location privacy preferences

Inspired by Westin (2003), we built a privacy model for location sharing from the study described in Section 2.1.

We performed simulations utilising this privacy model, in three different modes:

- *Default (D)*: Privacy preferences ignored.
- *Friendly (F)*: Privacy mode, where nodes may share to everyone, to no-one, or to their social contacts.
- *PubPriv (PP)*: Stronger privacy mode, where nodes share only with either everyone or no-one.

Figure 2 demonstrates our main finding. We found (Parris and Henderson 2011a) that users’ location-privacy preferences may significantly impact opportunistic network routing performance. Indeed, under the stronger *PubPriv* mode, the median delivery performance was zero.

2.3. Performance evaluation for protocols preserving social graph privacy

We target the threat of leakage of social-graph information by obfuscating the sender’s social network at the time of

³Note that Parris et al. (2010) describes only the first two runs, since the experiment was ongoing at the time of publication.

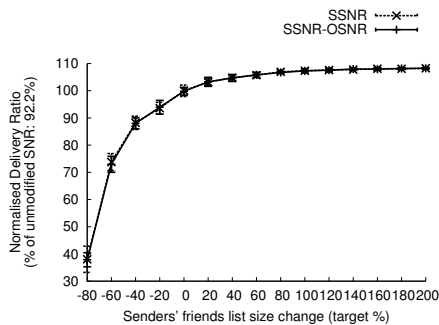


Figure 3: Delivery ratio after various degrees of obfuscation of social-network information. It is possible to maintain good performance (90% that of unmodified routing) after significant (−40%) obfuscation of the social network information.

message generation. Through simulation, we evaluate the performance of these new privacy-enhanced protocols.

Figure 3 demonstrates our main finding. It is possible to significantly obfuscate the social network information, by removing up to 40% of the sender’s “friends list” (social graph neighbours) from each message at generation time, while still maintaining good routing performance — a message delivery proportion (*delivery ratio*) of 90% that of unmodified social network routing). Further details are available in Parris and Henderson (2011b).

3. FUTURE WORK

We are planning an HCI experiment to probe the question of how privacy concerns may impact the willingness of users to participate in a future opportunistic network application. Our current results (Section 2.1) suggest that users behave differently in real and simulated systems, and thus we are planning a deceptive user study, where we simulate an opportunistic application, but do not inform participants that it is a simulated application. We intend to use the experiment results to inform interpretation of our current routing performance results (described in Sections 2.2 and 2.3).

3.1. Experiment plan

The opportunistic network application that we intend to simulate is a proposed, distributed, privacy-aware mobile advertising system, *MobiAd* (Haddadi et al. 2010). In such a real system, advertisements are distributed to mobile devices opportunistically, as are anonymous “click reports” describing interaction with the advertisements. The users, therefore, would dedicate some of their device’s resources to the network — for example, accepting a certain reduction in battery life, due to the extra energy used when transmitting messages.

By simulating this application, and presenting the experimental participants with various examples of potentially-sensitive information that may be leaked

through its use, we wish to investigate differing degrees of willingness to participate in the network, as measured by the quantity of phone resources that participants would be willing to dedicate to the application.

We hypothesise that, as we increase the amount of sensitive information displayed, users will become less willing to participate in the network — and, therefore, will allocate less resources to the application. This is a non-trivial hypothesis: perhaps the increased transparency of the application increases user confidence, and thus increases the willingness to participate in the network? We thus search for the amount of information “leakage” that results in maximal participation in the network.

In our proposed experiment, we will create an Android application, installed on mobile phones given to experiment participants for one week. This application will prompt participants for a username, on the first run. It will display (fake) adverts on the phone’s home screen at all times, by using Android’s capability to detect the user’s location, and then scraping web services for nearby businesses to “advertise” to the user.

There will additionally be an option on the home screen to display extra information about other participants. This information will be simulated for each participant, but the participants will not know this: this is the deception mentioned earlier, in order to mitigate the difference of behaviour of users of real vs simulated applications. The intention is that each participant will become aware that *their own* information may be leaked to other participants, by analogy to the information that they believe they can see about the others (“if I can see this information about them, then they must be able to see this information about me too”).

The simulated information displayed to the participants will differ across groups, to which participants will randomly be assigned at the start of the experiment. We intend to examine the effects of leakage of two types of potentially-sensitive information, at differing resolutions: *location* and *social graph*. We therefore propose dividing the participants into five groups:

1. *Control*: No sensitive information displayed.
2. *Location-Street*: Simulated locations of users displayed, at the street-level.
3. *Location-City*: Simulated locations of users displayed, at the city-level.
4. *Social-Neighbours*: Simulated list of participants frequently encountered by this particular user.
5. *Social-All*: Simulated list of all participants, with the option to click through to see their (simulated) lists of frequently-encountered participants.

The phones will prompt each participant once per day to answer a question, along with providing an explicit view of the simulated information (so that this information is salient, if the user chooses not to review it during the day). The question will be to ask the participant to use a slider (initially unset) to choose a maximum amount of battery life (in hours) that they would be prepared to devote to the opportunistic network application in the next 24 hours. We reason that this question would be easy for participants to understand, and quick to answer. The application would, however, not really use their phone's resources based on the previous question answers, so as to avoid draining the device's battery and possibly hindering the experiment: we reason that even without real usage of the battery, we would obtain a measure of the user's willingness to participate in the network at various times. We are interested only in *relative differences between groups*, rather than absolute values.

As mentioned earlier, the experiment results would inform interpretation of our previous performance results, which evaluate performance for protocols under varying assumptions of user behaviour. Example questions that we hope to answer are: (i) Which privacy mode from our location simulation is more realistic? (ii) Are participants concerned about social graph privacy?

3.2. Problems and questions

Some open questions, which we hope to discuss during the HCI 2011 Doctoral Consortium, include:

- How precisely could we best display potentially-large numbers of simulated encounters (i.e., social graph information) or location information to the relevant participants, on the small phone screens?
- Should participants interact to differing degrees with the visualisations of potentially-sensitive information, then what are the implications for interpretation of the results?
- How could the potentially-noisy results best be interpreted?
- How many participants should take part in the experiment? And how should they be selected?
- If participants' willingness to participate in the network varies in different contexts (e.g., perhaps there are particular times when they do not wish to give up any battery life), then is it possible to improve on the coarse slider-based questions?

4. CONCLUSION

We have summarised our current results, and detailed our planned user study. We hope that these results and plan may be of interest to HCI 2011 Doctoral Consortium participants, and in turn we hope to receive useful feedback and suggestions.

5. ACKNOWLEDGEMENTS

The authors thank Hamed Haddadi and Pan Hui for their ideas and suggestions regarding the planned MobiAd experiment.

6. REFERENCES

- C. Boldrini, M. Conti, and A. Passarella. Exploiting users' social relations to forward data in opportunistic networks: The HiBOP solution. *Pervasive and Mobile Computing*, 4(5):633–657, October 2008. ISSN 15741192. doi: 10.1016/j.pmcj.2008.04.003.
- S. Consolvo and M. Walker. Using the experience sampling method to evaluate ubicomp applications. *Pervasive Computing, IEEE*, 2(2):24–31, 2003. doi: 10.1109/MPRV.2003.1203750.
- E. M. Daly and M. Haahr. Social network analysis for information flow in disconnected delay-tolerant MANETs. *IEEE Transactions on Mobile Computing*, 8(5):606–621, May 2009. ISSN 1536-1233. doi: 10.1109/TMC.2008.161.
- H. Haddadi, P. Hui, and I. Brown. MobiAd: private and scalable mobile advertising. In *Proc. MobiArch 2010*, pages 33–38, New York, NY, USA, September 2010. ACM. ISBN 978-1-4503-0143-5. doi: 10.1145/1859983.1859993.
- P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. In *Proc. MobiHoc 2008*, pages 241–250, New York, NY, USA, May 2008. ACM. ISBN 978-1-60558-073-9. doi: 10.1145/1374618.1374652.
- I. Parris and T. Henderson. The impact of location privacy on opportunistic networks. Under submission, February 2011a.
- I. Parris and T. Henderson. Privacy-enhanced social-network routing. *Computer Communications*, 2011b. ISSN 01403664. doi: 10.1016/j.comcom.2010.11.003. In press.
- I. Parris, F. Ben Abdesslem, and T. Henderson. Facebook or Fakebook?: The effect of simulation on location privacy user studies. In *Proc. PUMP 2010*, London, UK, September 2010. BCS. URL <http://scone.cs.st-andrews.ac.uk/pump2010/papers/parris.pdf>.
- L. Pelusi, A. Passarella, and M. Conti. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *IEEE Communications Magazine*, 44(11):134–141, November 2006. ISSN 0163-6804. doi: 10.1109/MCOM.2006.248176.
- Alan F. Westin. Social and political dimensions of privacy. *Journal of Social Issues*, 59(2):431–453, July 2003. doi: 10.1111/1540-4560.00072.