Performance Evaluation of Delay-Tolerant Wireless Friend-to-Friend Networks for Undetectable Communication

Ana Barroso
Technische Universität Darmstadt
Secure Mobile Networking Lab
Mornewegstr. 32, 64289 Darmstadt, Germany
Email: abarroso@seemoo.tu-darmstadt.de

Matthias Hollick
Technische Universität Darmstadt
Secure Mobile Networking Lab
Mornewegstr. 32, 64289 Darmstadt, Germany
Email: mhollick@seemoo.tu-darmstadt.de

***Abstract*—Anonymous communication systems have recently increased in popularity in wired networks, but there are no adequate equivalent systems for wireless networks under strong surveillance. In this work we evaluate the performance of delay-tolerant friend-to-friend networking, which can allow anonymous communication in a wireless medium under strong surveillance by relying on trust relationships between the network's users. Since strong anonymity properties incur in performance penalties, a good understanding of performance under various conditions is crucial for the successful deployment of such a system. We simulate a delay-tolerant friend-to-friend network in several scenarios using real-world mobility data, analyze the trade-offs of network-related parameters and offer a preliminary throughput estimation.**

Keywords: wireless, delay-tolerant, anonymous communication, undetectability, friend-to-friend networks

I. Introduction

To allow for undetectable wireless communication among mobile devices, ad-hoc networks may prove essential to avoid infrastructure that can be censored or monitored by adversaries for the purposes of identification. State-level adversaries can identify mobile device owners based on the information from their network carrier. Moreover, Montjoye et al. were able to identify 95% of 1.5 million people based on anonymized location data of their mobile phones [5].

However, avoiding infrastructure is not the only step required to evade surveillance. The wireless medium is, due to its nature, easy to monitor by a local adversary. We need a mechanism that hides the paths of the messages as they travel from the sender to the destination. Systems such as anonymous remailers [7] or Tor [6] can increase the privacy of their users on wired networks, but they can hardly be transposed to a scenario of moving wireless nodes, where the high mobility implies constantly changing link states. A precomputed sequence of proxy nodes is not feasible as a routing path in such a network.

aDTN, the wireless network scheme we propose in [3], relies on friend-to-friend communication and on the delays introduced by the store-carry-and-forward approach of delay-tolerant communication to provide strong anonymity in wireless scenarios. It ensures undetectability[1] of messages, as well as of the acts of sending and receiving them, and therefore hiding the relationships between users. However, an analysis of its network performance has not been carried out so far, and is essential to understand its limitations.

In this work we analyze the performance of wireless friend-to-friend delay-tolerant networking, the underlying network model of *aDTN*. Since the performance of the network is highly dependent on the contacts between friends, we use real-world mobility data to simulate the network and measure its performance under various conditions.

This work is structured as follows: we start by explaining the network model in Section II. In Section III we analyze the network performance after introducing the simulation procedure and performance metrics. We also estimate throughput based on the payload size of *aDTN*. In Section IV we describe related work. Finally, in Section V we conclude our work and outline our next steps.

II. Network and communication model

A friend-to-friend network is a particular case of a private peer-to-peer network [14] where nodes only send packets to other nodes they know and trust. The scheme requires that the nodes have a previously established trust relationship. In our model this relationship has only two states: either two nodes trust each other, or they do not. If they do, they are said to be *friends*. A trust relationship can be extended to a larger number of nodes: we speak of a group of friends if all nodes in the group trust every single other member. For simplicity, when we mention a group throughout this work, we mean a group of friends.

The network consists only of mobile nodes which wirelessly transmit packets at a constant interval, independently of whether other nodes are in range or not. Each packet transmitted by a node contains a message that can be read only by the members of one of the groups the node belongs

[1] In this work we use the privacy terminology described by Pfitzmann and Hansen in [12].

978-1-4673-6770-7/15/$31.00 ©2015 IEEE

to. Every node in range receives the packet and attempts to read the message contained in it. Messages that cannot be read are discarded. When a node receives a message that it can read, the message is stored for later retransmission, together with other messages stored by the node. To avoid linking messages to the users, the packets are cryptographically transformed so that sending the same message multiple times will result in different-looking packets that cannot be correlated by an observer. Nodes may belong to multiple groups, which allows messages from a group to spread to other groups.

The aim of the network scheme is to spread the messages through the network in a best-effort fashion, and duplicates are allowed for this effect. While this is expensive from a network point of view, it is required by *aDTN* for achieving communication undetectability. The protection goals and effectiveness of the scheme are explained in detail in [3].

III. Evaluation of network performance

While gossiping-based message dissemination has been extensively studied before [4], [8], [9], [11], there is little work in friend-to-friend delay-tolerant networking, which imposes new restrictions. The goal of this evaluation is to understand under which conditions friend-to-friend delay-tolerant networking is viable in a scenario of mobile wireless nodes. In this section we describe the simulation procedure and the relevant metrics to our analysis. Finally we analyze the simulation results under various parameters and estimate throughput of the *aDTN* scheme based on its packet size.

To characterize network performance we selected the following metrics:

- *sent* is the number of sent packets
- *heard* is the number of sent packets received by at least one node
- *unheard* is the number of packets transmitted by a node when no other node was in its range
- *copies* is the number of all received copies of a sent packet
- *readable* is the number of copies where the node was able to read the message contained in the packet, i.e. the receiver was in the same group as the sender
- *unreadable* is the number of copies where the node could not read the message from the packet and discards it, i.e. the receiver was not in the same group as the sender.

We conduct simulations of our scheme in a mobility scenario based on real-world data obtained from the Nokia Mobile Data Challenge [10]. This dataset contains the GPS traces of 186 smartphone users over a period of 1.5 years in the region surrounding the city of Lausanne. We use a 2-day sample of this data restricted to the 50 most active nodes in that period. Using the BonnMotion [1] mobility framework we infer the time intervals when pairs of nodes were within 15m of each other. These time intervals we call contacts, and the times between meetings are called inter-contact times.

The simulation[2] runs as follows: whenever a node is scheduled to transmit, we determine which nodes are in range. In case a node is in range, we verify if it belongs to the same group as the packet sender. If so, that represents a readable message and the receiver keeps it; if not, it represents an unreadable message and the receiver discards the packet.

[2]Our simulator code is free software, available online at http://www.seemoo.tu-darmstadt.de/research/software/f2fdtnsim/

A. Effect of transmission interval on received packets

A node's transmission frequency is defined in our experiment by the interval *i* between sending two packets. A smaller interval leads to more data transmitted, which can be desired in scenarios that need high throughput. However this can impact the processing capacity of the devices, or their battery power, so it may be necessary to set it only as low as the application scenario requires it. The interval can be arbitrarily varied, so we selected *i* = 30, 60, 120, 240, measured in seconds, to obtain a first impression of its effect.

In Table I we show the impact of various interval lengths on the transmission and reception of packets. *%heard* and *%copies* are relative to the number of transmissions. We can observe that the rates of heard and received packets remain approximately constant over the range of intervals. This is due to the fact that, in the used dataset sample, nodes establish contacts that span much longer than the interval, which increases the chance of a transmission to fall into the time window of a contact. We expect that in scenarios with short contacts, increasing the interval significantly impacts the network throughput.

The number of copies represents the maximum number of opportunities to forward data in the network. Their success depends on group membership, which we analyze next.

B. Effect of group characteristics on delivered messages

Given the contacts in the dataset sample, we generated groups based on the graph $G = (E, V)$, where the vertices V are the network nodes and the edges E link all the pairs of vertices (v_1, v_2), $v_1, v_2 \in V$ which have met at least once. Meeting nodes are candidates to represent friends and thus eligible to build groups. Hence, from the cliques of G we generated groups with at most *max_group_size* vertices, and where each vertex v is in at most *group_limit* groups.

According to [3], *group_limit* and *max_group_size* should be kept low for security purposes, hence we selected the following parameter values:

- *group_limit*: 2, 3, 4, 5, 6
- *max_group_size*: 2, 3, 4, 5, 6

Additionally, we ran the simulation on a baseline scenario where all nodes belong to the same group (and therefore all messages are readable by their receivers).

Figure 1 compares the average number of readable copies across all simulation runs. We notice that some combinations of *max_group_size* and *group_limit* have better results than the rest. This is explained by the random distribution of nodes in the groups that assigned a few nodes that meet each other often to the same group. While this is an inconvenience to our analysis, we can also get an insight into the advantage of forming groups with trusted people, which are more likely

TABLE I: Effect of transmission interval i on the number of received packets and throughput

i	sent	unheard	heard	copies	%heard	%copies	sending rate	received goodput
30s	7925	5141	2784	13941	35.1%	175.9%	533.6bps	942.1bps
60s	4044	2718	1326	6639	32.8%	164.2%	272.0bps	451.0bps
120s	2028	1333	695	3482	34.2%	171.7%	136.8bps	235.4bps
240s	1025	678	347	1742	33.9%	167.0%	68.8bps	117.8bps

(a) Avg. readable copies for i = 30s (b) Avg. readable copies for i = 60s (c) Avg. readable copies for i = 120s (d) Avg. readable copies for i = 240s

Fig. 1: Average number of readable copies, across all parameter combinations.

to meet at regular intervals, in contrast to unknown people. We see that in the scenarios where friend nodes contact each other regularly, the results outperform the other scenarios by a factor of 4 to 8.

If we abstract from those extreme results, however, we can detect a slight trend in higher group sizes leading to higher numbers of readable copies. This is more visible in Figures (a) and (b), where the color darkens towards the right. It is indeed to be expected that if the maximum group size increases, the number of readable copies increases as well, since the likelihood to transmit a packet to a node of the same group is higher. On the other hand, a too high number of groups may decrease it, since a node has more groups to distribute the messages to, and each packet can only be read by one group. Therefore, the likelihood that the transmitted packet contains a message readable by a friend in range diminishes. This can be compensated by raising the maximum group size. However, the impact of group size and number of groups is negligible in comparison to building optimal groups in terms of meeting frequency.

C. Estimation of network throughput

The number of message copies corresponds to the network throughput. If we multiply the number of packets sent by a node by the maximum payload size and divide by the simulation duration, we obtain the payload data rate, also called goodput. We use the definition of goodput offered in IETF's RFC5166[3]: "measured in bytes per second, [it] is the subset of throughput consisting of useful traffic".

[3]http://tools.ietf.org/html/rfc5166

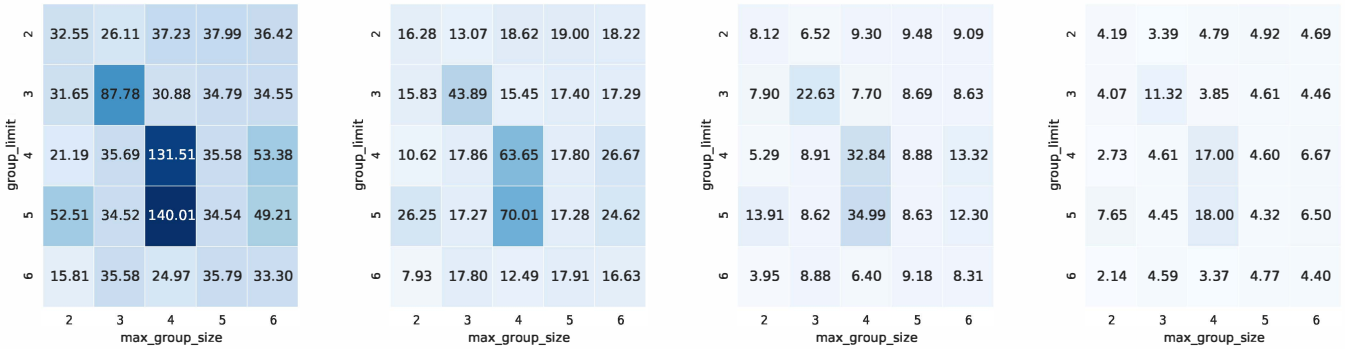
Since the number of packets sent is different from the number of received packets due to the copies, we define the former as sending rate and the later as received goodput. While in RFC5166 the term goodput excludes duplicate packets, in this work we consider it to include the copies, as they are useful traffic for the present network scheme.

In Table I we compare the sending rate with the baseline received goodput, i.e. received goodput in the best-case scenario where all nodes are in the same group and, hence, all messages are readable by the destination. As an example we use the maximum payload size of 1454 bytes used by *aDTN*. Since the received goodput is directly proportional to the number of readable copies, it varies with the number of groups and their size. In Figure 2 we show the corresponding calculation of received goodput values for all combinations of *group_limit*, *max_group_size* and transmission interval i.

While these goodput values are quite low for today's communication standards, a transmission every 30s is orders of magnitude below the network utilization rate of a typical wireless LAN link. We did not simulate shorter intervals for practical reasons, but if the trend shown by our data holds, at 1s transmission interval a node can reach on average a goodput of approximately 30Kbps in theory. In practice, we expect lower throughput due to collisions, transmission errors and other physical layer aspects. A thorough practical analysis in real-world network conditions is left for future work.

IV. Related work

In the friend-to-friend Turtle network [13] nodes only address packets to friendly nodes. This system is an overlay to the IP network, where the problems of the wireless medium and of delay-tolerant communication are not present.



(a) Average goodput for i = 30s (b) Average goodput for i = 60s (c) Average goodput for i = 120 (d) Average goodput for i = 240s

Fig. 2: Estimated average received goodput per node, measured in bytes per second, for an example packet size of 1488 bytes.

Thilakarathna et al. [15] propose friend-to-friend opportunistic content dissemination to improve trust, delivery latency and privacy issues in mobile social networks. It also requires previous trust establishment between users, but unlike the scheme analyzed in this work, it needs a central entity that analyzes content access patterns of users to predict which content they might be interested in. Based on those predictions, the central entity selects users, called helpers, that are more likely to carry the content to the users that are potentially interested in it. Another difference is that, in their system, nodes do not need to transmit packets at regular intervals due to different protection goals.

Another related system is HumaNets [2], a routing protocol for decentralized delay-tolerant networks that exploits recurring mobility patterns to improve performance. HumaNets does not impose such a strict restriction on links as *aDTN* does, since its protection goal is limited to location privacy.

V. Conclusion

In this work we study the network performance of wireless friend-to-friend delay-tolerant networking. We simulate the network behavior of the *aDTN* [3] scheme using real-world mobility data and show the trade-off between throughput and the number and size of friend groups, as well as transmission frequency.

We show that increasing transmission frequency proportionally increases throughput. Since devices are physically limited in this regard, another way to further improve throughput is to increase the maximum size of a friend group. However, this solution is not viable for the *aDTN* scheme for security reasons, but it may be acceptable in other wireless friend-to-friend networks based on trust groups.

Additionally, we observed that the social graph of network members plays a decisive role in throughput, overshadowing parameters such as transmission frequency and number of trust groups and their size.

The obtained results give us an insight into the feasibility of wireless friend-to-friend protocols, but also opened new questions. Our next steps are: a) to better understand the impact of the structure of the social graph on throughput; b) to implement the *aDTN* scheme and measure its performance in a real-world scenario; and c) to analyze message dissemination and the effects of malicious network activities such as jamming, spamming and blackhole attacks.

References

[1] N. Aschenbruck, R. Ernst, E. Gerhards-Padilla, and M. Schwamborn, "BonnMotion: a mobility scenario generation and analysis tool," in *ICST'10*, 2010, pp. 51:1–51:10.

[2] A. Aviv, M. Sherr, M. Blaze, and J. Smith, "Privacy-aware message exchanges for geographically routed human movement networks," in *Computer Security ESORICS'12*, 2012, vol. 7459, pp. 181–198.

[3] A. Barroso, "aDTN - undetectable communication in wireless delay-tolerant networks," 2014. [Online]. Available: https://www.seemoo.tu-darmstadt.de/dl/seemoo/seemoo-tr-2015-01.pdf

[4] J.-C. Bermond, L. Gargano, A. A. Rescigno, and U. Vaccaro, "Fast gossiping by short messages," *SIAM Journal on Computing*, vol. 27, no. 4, pp. 917–941, 1998.

[5] Y.-A. de Montjoye, C. Hidalgo, M. Verleysen, and V. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," *Scientific reports*, vol. 3, 2013.

[6] R. Dingledine, N. Mathewson, and P. Syverson, "Tor: The second-generation onion router," DTIC Document, Tech. Rep., 2004.

[7] M. Edman and B. Yener, "On anonymity in an electronic society: A survey of anonymous communication systems," *ACM Computing Surveys (CSUR)*, vol. 42, no. 1, p. 5, 2009.

[8] S. M. Hedetniemi, S. T. Hedetniemi, and A. L. Liestman, "A survey of gossiping and broadcasting in communication networks," *Networks*, vol. 18, no. 4, pp. 319–349, 1988.

[9] S. Khuller, Y.-A. Kim, and Y.-C. J. Wan, "On generalized gossiping and broadcasting," *J. Algorithms*, vol. 59, no. 2, pp. 81–106, May 2006.

[10] J. K. Laurila, D. Gatica-Perez, I. Aad, O. Bornet, T.-M.-T. Do, O. Dousse, J. Eberle, M. Miettinen *et al.*, "The mobile data challenge: Big data for mobile computing research," in *Pervasive Computing*, 2012.

[11] M.-J. Lin, K. Marzullo, and S. Masini, "Gossip versus deterministically constrained flooding on small networks," in *Distributed Computing*, M. Herlihy, Ed., 2000, vol. 1914, pp. 253–267.

[12] A. Pfitzmann and M. Hansen, "A terminology for talking about privacy by data minimization: Anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management," *Version 0.34 Aug*, vol. 10, 2010.

[13] B. C. Popescu, B. Crispo, and A. S. Tanenbaum, "Safe and private data sharing with Turtle: friends team-up and beat the system," in *Security Protocols*. Springer, 2004.

[14] M. Rogers and S. Bhatti, "How to disappear completely: A survey of private peer-to-peer networks," *Networks*, vol. 13, p. 14, 2007.

[15] K. Thilakarathna, A. C. Viana, A. Seneviratne, and H. Petander, "Mobile social networking through friend-to-friend opportunistic content dissemination," in *Proceedings of the 14th ACM MobiHoc*, 2013, pp. 263–266.