# Are You Moved by Your Social Network Application?

Abderrahmen Mtibaa*        Augustin Chaintreau*        Jason LeBrun†

Earl Oliver‡        Anna-Kaisa Pietilainen*        Christophe Diot*

Thomson
46 quai A. Le Gallo
92648 Boulogne Cedex
France

## ABSTRACT

This paper studies a Bluetooth-based mobile social network application deployed among a group of 28 participants collected during a computer communication conference. We compare the *social graph* containing friends, as defined by participants, to the *contact graph*, that is the temporal network created by opportunistic contacts as owners of devices move and come into communication range. Our contribution is twofold: first, we prove that most properties of nodes, links, and paths correlate among the social and contact graphs. Second, we describe how the structure of the social graph helps build forwarding paths in the contact graph, allowing two nodes to communicate over time using opportunistic contacts and intermediate nodes. Efficient paths can be built using only pairs of nodes that are socially close (*i.e.* connected through a few pairs of friends). Our results indicate that opportunistic forwarding complies with the requirement of social network application.

## Categories and Subject Descriptors

C.2.1 [**Network Architecture and Design**]: Network topology; C.2.0 [**General**]: Data communications

## General Terms

Measurement, Performance

## Keywords

Online social network, mobile network, delay-tolerant network, pocket-switched networks, temporal networks, centrality

---

*email: *firstname.lastname@thomson.net*
†email: *jblebrun@ece.ucdavis.edu*
‡email: *eaoliver@uwaterloo.ca*

## 1. INTRODUCTION

Social life primarily takes place through physical meeting. Recently, however, the telegraph and telephone networks made a first step toward remote social interaction. More recently, the Internet added multiple social interaction techniques not based on physical meeting: email, chat, and Online Social Network services (OSN) such as Facebook, Orkut, MySpace, and LinkedIn. These applications create a virtual space where users can build their social network independently of where they are located, and allow these social networks (or communities) to interact freely using a large set of Internet applications. However, when people with similar interests or common acquaintances get close to each others in streets or conferences, they have no automated way to identify this potential "relationship". With geolocalization applications, it is now highly likely that OSNs will include in a near future some representation of user location, and offer services to "link" mobile users. However, the relation between virtual social interactions and physical meeting remains largely unexplored.

In this paper we study the evolution of the social relationships of a group of 28 participants using smartphones containing a mobile opportunistic social networking application. Our study was conducted at the ACM CoNEXT conference in New York, December 2007. At the beginning of the conference, the application asked users to select their friends among conference participants. We call this list of social connections the *initial social network*. During the conference, experiment participants are informed of other participants within Bluetooth range, or *neighborhood*. Participants are notified by a distinctive ring or a vibration if a friend or a friend of a friend is detected. A user can decide to ignore these notifications, or meet one of these persons physically and/or add them to its list of friends. Our devices log all Bluetooth contacts between experimental devices, as well as all user action such as adding a new friend or deleting an existing friend. We use this data set to study the evolution of the initial social graph and to analyze how human mobility and social relationships mutually impact each other in the specific context of a conference. We have decided to conduct these experiments during conferences as it is a reasonable group size to cover and because it can be reproduced with similar environment. As human mobility creates opportunistic contacts, these could be used to deliver data between two nodes, on a forwarding path using intermediate nodes (see, for instance, Pocket Switched Net-

works [1] or Delay-Tolerant Networks, `www.dtnrg.org`). We study how such opportunistic forwarding can benefit from social properties.

To the best of our knowledge, this is the first attempt to compare the graph of social relationships as defined by the users and the contact opportunities resulting from their mobility during a community event (*i.e.* a computer network conference). We make the following contributions:

- Properties of nodes, links, and paths, are studied jointly in the social graph of friends and in the temporal network of opportunistic contacts. We observe expected and unexpected similarities, which confirm that classifying nodes based on their friends is relevant for the temporal network as well. We notice that this correspondence improves with time. (Section 4)

- We show that delay-efficient forwarding paths can be constructed using only contacts between people close in a social sense, following the principle underlying most OSN. Moreover, the contacts that are critical for opportunistic forwarding can be identified from the position of the contacted nodes in the social network. Our empirical analysis compares several heuristic rules. (Section 5)

These early results, even if limited in scope and depth, are encouraging and will help us improve our experimental software in preparation for future experimental campaigns.

## 2. RELATED WORK

Most social properties have been studied for static graph, with two notable exceptions [7, 9]. Properties of paths built over time in a quickly varying graph is a relative new topic [9, 2, 8]. So far the similarities with traditional social networks have been investigated in a macroscopic sense: evidence of heavy tailed statistics for degree [8] and inter-contact times [2, 5], community identification [9, 4], short diameter [7, 2]. In contrast, here we compare the topological properties of each node as observed in two related networks.

Recent work has considered the problem of designing opportunistic forwarding schemes that are aware of social properties [3, 4]. This work implicitly assumes that opportunistic contacts relate with the social property that is used to design one algorithm. Our work does not propose new algorithms; it addresses the above issue more generally. It can be used to understand what type of information is the most relevant. Note that for the first time, social relationships between participants is defined by the participants using the application. Previous experiments [4] have asked participants their interests, affiliation, etc. and social graphs have been inferred from their answers.

## 3. DATA AND METHODOLOGY

### 3.1 Experimental settings

The goal of the experiment is to study the characteristics of a social networks built using mobile devices. Users may meet potential friends, make new friends or delete existing friends. The experimental devices are HTC s620 and Touch smartphones running our mobile opportunistic social networking application.

Participants have been chosen among the attendees of the ACM CoNext 2007 conference held in New-York City on December, 2007. Before running the application, each participant is asked to select its friends among the 150 CoNext participants. The list of friends constitutes the initial social network of each participant. Our social networking application rings or vibrates any time a friend, a friend of a friend[1], comes within Bluetooth range (approximately ten meters). The neighborhood is displayed on the user's device who can add new friends or delete existing friends based on his or her discussion with them. Our objective is to study how the initial social network relates to opportunistic contacts made during the conference and how this network evolves.

Limiting the list of friends to the set of conference participants is not representative of their complete social network, nevertheless it captures social relationship within this event. We believe it is an acceptable restriction for an initial experiment.

The 28 participants were asked to use the application at their own convenience and to maintain the device with enough battery. The experiment lasts three days and each device is being used an average of 2.2 days (as people arrive and leave at different times)[2]. We define the offline time as the time during which (within the full trace) the device was not running the application. The average offline time is 32.1 hours resulting mainly from application crashes and battery depletion. Devices scan their environment using Bluetooth every 2 minutes (as in previous experiments measuring opportunistic contacts [1]). A total of 9024 opportunistic contacts have been recorded among the experimental devices. For this particular experiment, we ignore contacts made with non experimental Bluetooth devices such as laptops or regular cell phones. We focus in the rest of the paper on day-time contact characteristics (contacts occurring during conference time: from 9am to 6pm), to avoids the long disconnections occuring because of evening and night time.

### 3.2 Definitions and Terminology

In the rest of this paper, we refer to the slowly varying graph of friendship between participants as the *social graph*, which we denote as $G = (V, E)$.

The collection of opportunistic Bluetooth contacts between the participants form a temporal network (*i.e.* a graph with a static set of nodes, and a set of edges that may change with time, see [6, 1]). We call this temporal network the *contact graph*, and denote it by $G_t = (V, E_t)$. Paths may be constructed in temporal network as a concatenation of contacts following a chronological property. Among these paths, a path from $s \in V$ to $d \in V$ starting at time $t_0$ is delay-optimal if it reaches the destination $d$ in the earliest possible time. Delay-optimal paths for any starting time and any source-destination pair can be computed efficiently via dynamic programming (see [2] for more details).

## 4. TOPOLOGICAL COMPARISON

In this section, we study the topologies of the social graph and the contact graph with respect to nodes, contacts and paths.

We start by some general description of the social graph: Table 1 presents some characteristics of the social graph cre-

---

[1] A friend of a friend is detected by finding a common element in the nodes's lists of friends. This computation may raise some privacy concerns, we ignore that point here.
[2] One out of the 28 nodes has been removed as it was used only for a couple of hours.

ated by participants in the application described above, it is shown in Figure 1.

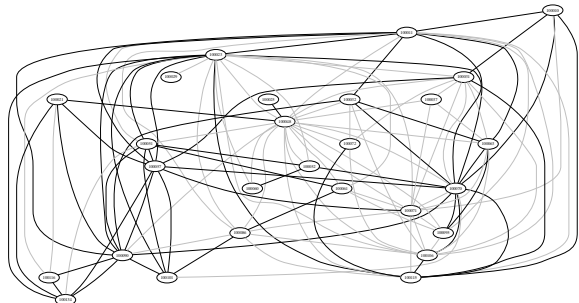|                        | Initial Graph | Final Graph |
|------------------------|:-------------:|:-----------:|
| # nodes                | 27            | 27          |
| largest conn. comp.    | 26            | 27          |
| # edges                | 68            | 129         |
| average degree         | 5.2           | 9.5         |
| clustering coefficient | 0.2           | 0.36        |
| diameter               | 7             | 4           |

**Table 1: Statistics of social graph.**



**Figure 1: Initial social graph and final social graph of friends among participants of CoNEXT experiment.**

The *initial social graph* (represented by black edges in Figure 1) denotes the friends network when the application starts. During the experiment users could become friends when meeting opportunistically, resulting in the *final social graph* which contains as well the gray edges shown in the same figure. From the table, we observe that the average degree and the clustering coefficient[3] roughly doubles during the course of the experiment. The diameter of the social graph decreases by a factor of two.

## 4.1 Properties of node

Complex system and social networks are usually characterized by large heterogeneity between the nodes. A small portion of highly active nodes typically co-exist with a large population of nodes that follow a normal volume of activity. In this section, we wish to identify such nodes and compare them in both graphs.

### 4.1.1 Node degree

A first estimation of the importance of a node in a social network is its *degree*, or its number of friends, which measures the size of its neighborhood in the social graph. Similarly, in the contact graph, the size of the neighborhood of a node is the average number of devices that it meets via opportunistic contacts during a scan. Figure 2 compares the two measures of degree (in the social graph and the contact graph) jointly for each node.

We expect that a node with a larger number of friends also sees more opportunistic contacts, although these two values

---

[3]The clustering coefficient is defined as the average for $u \in V$ of the ratio between the number of edges connecting neighbors of $u$ and the number of pairs of neighbors of $u$.
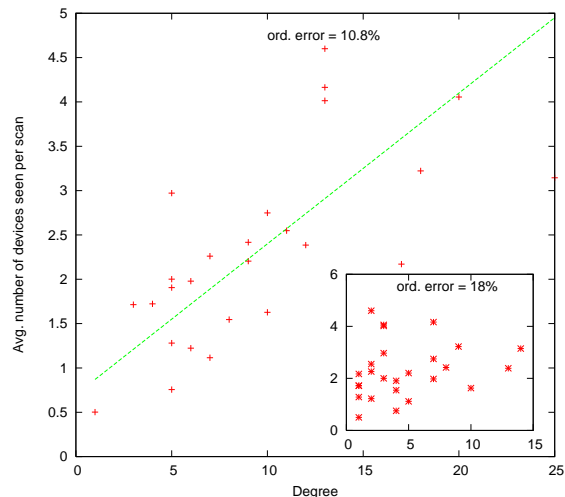


**Figure 2: Joint values of the nodes degree in the social and contact graph: for final social graph (main) and initial social graph (small frame).**

may not necessarily be related linearly. Hence, to check the first assertion, we define the *ordering error* as:

$$\frac{|\{ \ u \neq v \in V \,, \ M(u) \leq M(v) \ \text{and} \ M_t(u) > M_t(v) \ \}|}{|V|(|V|-1)},$$

where $M$ and $M_t$ denotes in general two functions $V \to \mathbb{R}$, which depends respectively on $G$ and $G_t$, and associates a metric value to each vertex. For the two degree metrics already defined above, we indicate the ordering error on the figure.

There is not a significant correlation among the two metrics when the contact graph is compared to the initial social graph (indeed, a fifth of the pairs create ordering errors when using these two metrics). However, during the course of the experiment, some correlation appears and the ordering error decreases. In other words, it is likely that we meet regularly people who are not part of our social circle, but as time goes on, the proximity plays a role in order to make friends. This applies to our experiment, since our application allows participants to make new friends only via opportunistic meeting.

### 4.1.2 Centrality

Centrality is a more refined measure of the importance of a node for a network; it deals with the occurrence of this node inside the shortest paths connecting pairs of other nodes. Centrality of a node is defined in a static graph as:

$$C(v) = \frac{|\{ \ s, d \in V \setminus \{v\} \,, \ s \neq d \ | \ v \in \pi(s,d) \ \}|}{(|V|-1)(|V|-2)},$$

where $\pi(s,d)$ denotes the shortest path from $s$ to $d$ in $G$. Similarly, we define the centrality of a vertex in the temporal network $G_t$ as

$$C_t(v) = \frac{|\{ \ t, s, d \in V \setminus \{v\} \,, \ s \neq d \ | \ v \in \pi(s,d,t) \ \}|}{(|V|-1)(|V|-2) \cdot T},$$

where $\pi(s,d,t)$ denotes the delay-optimal path starting at time $t$ from $s$ to $d$, and $T$ is the experiment duration.

Figure 3 compares the centrality of nodes, shown jointly for social and contact graph, when initial and final social

graph are used. Centrality varies among nodes: one node appears in almost 7% (respectively, 3%) of the shortest paths drawn in the final social graph (respectively, contact graph) while most other nodes appear in less than 1% (respectively, 0.5%) of them. For the final social graph, the two measures correlate (and 96% of the pairs of node compare in the same way according to both measures). In the rest of this paper we will focus on centrality of nodes measured in the final social graph.
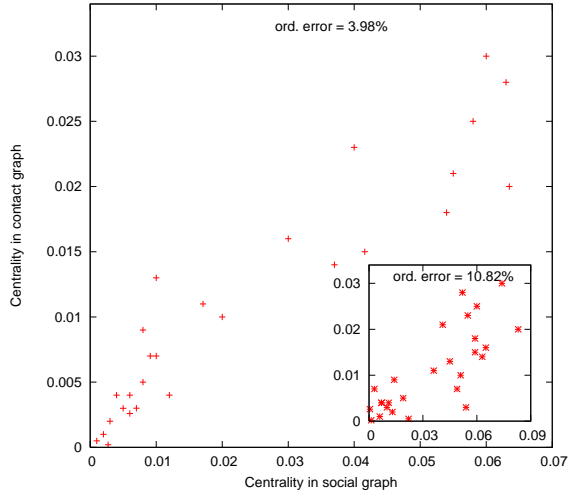


Figure 3: Joint values of the nodes centrality in the social and contact graph: for final social graph (main), initial social graph (small frame).

Centrality measures includes properties of multi-hop path, it is less affected by limitations of Bluetooth than the degree and should be a more accurate measuremenment. Our results indicate that beyond local discrepancy, hierarchical relations between the nodes should have profound relation in the two graphs.

## 4.2 Properties of contacts

We now study how properties of opportunistic contacts depends on the *social distance*, defined between two nodes on the social graph. We define friends as having a distance of one, friends of friends have a distance of two, and so on.

We have studied contacts according to their duration, their frequency, and the time elapsed between two successive contacts of the same pair (also known as *inter-contact time*). Due to space constraints, we present only the values of the inter-contact time in Figure 4. The median inter-contact time grows from 6 minutes between two friends, to nearly an hour (ten times more) when nodes have distance three or four in the social graph. We see that contacts between friends are almost all separated by less than an hour. We observe that the centrality of a nodes also plays an important role. A node with centrality around 6% has on average a contact every hour with every other node, independently of the social distance between them. Nodes with medium or low centrality (around 2% or 0.5%) see on average a contact every hour for each of their friends, but not with the nodes at distance two and more from them.

Finally, we observe that contacts between friends are sig-
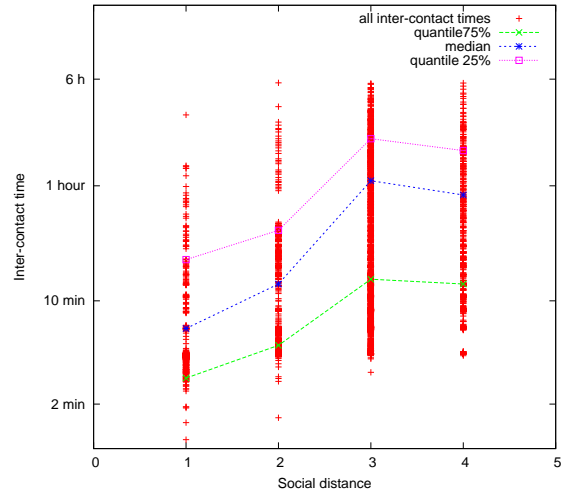


Figure 4: Inter-contact time seen for pairs with different distance in the social graph.

nificantly longer: 75% of the contacts with friends are longer than ten minutes. Whereas 75% of the contacts with nodes at distance four are shorter than thirteen minutes.

## 4.3 Properties of Delay-Optimal Paths

We now study delay-optimal paths (already defined in Section 3) as a function of the distance between the source and the destination in the social graph.

Figure 5 (a) plots the CDF for the optimal delay seen at all starting times, for sources and destinations with different social distance. As expected, delay is smaller for nodes that are closer. Note that this distribution depends significantly on the centrality of the source, as shown for two different nodes in Figures 5 (b) and (c). The delay from a central nodes to any other node is within ten minutes roughly 20% of the time, even for nodes at distance three. In constrast, the delay falls under ten minutes only for 10% of the time when the source is non-central, even to communicate with its friend.

The $(1 - \varepsilon)$-*diameter* of a temporal network is defined in [2]; it is the number of hops that is needed to achieve a fraction at least $(1 - \varepsilon)$ of the success ratio obtained with flooding, for a given maximum delay. In other words, this is the maximum number of hops $k$, necessary to construct all paths needed to be almost competitive with flooding, with a fixed maximum delay. Figure 5 plots the diameter (choosing $\varepsilon = 1\%$) for pairs of nodes at different social distance. Interestingly, the value of the diameter is of the same order as the diameter of the social graph. The diameter follows the social distance between the source and the destination: each additional hop increases the value of the diameter by one.

To summarize our results, nodes may be ranked according to their centrality, their rank in the social graph and in the contact graph coincide more or less. Opportunistic contacts, and the optimal paths which may be constructed between two nodes, depend on their centrality and on the social distance between them.
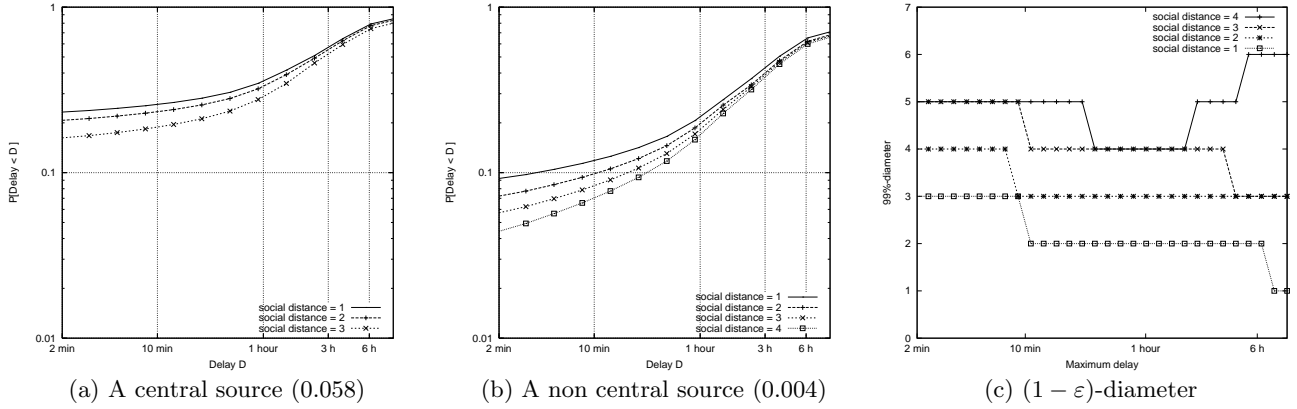
(a) A central source (0.058)  (b) A non central source (0.004)  (c) $(1 - \varepsilon)$-diameter

**Figure 5: Properties of delay-optimal paths for different distance in the social graph.**

# 5. PATHS CONSTRUCTION WITH OSN

The observations we made in the previous section can be used to design heuristic rules to construct efficient paths based on relationship in the social network. So far we have studied the properties of delay-optimal paths in the temporal network. Such paths offer the best possible delivery ratio, but they can only be found *a posteriori* or by using flooding. In this section, we present early results on the constructive properties of forwarding paths.

All the construction rules we consider fits in the following general model: depending on the source $s$ and the destination $d$, a rule defines a subset of directed pairs of nodes $(u \rightarrow v)$ so that only the contacts occurring for pairs in the subset are allowed in forwarding path. We consider the following construction rules.

**neighbor($k$):** $(u \rightarrow v)$ is allowed if and only if $u$ and $v$ are within distance $k$ in the social graph.

**destination-neighbor($k$):** $(u \rightarrow v)$ is allowed if and only if $v$ is within distance $k$ of $d$.

**non-decreasing-centrality:** $(u \rightarrow v)$ is allowed if and only if $C(u) \leq C(v)$.

**non-increasing-distance:** $(u \rightarrow v)$ is allowed if and only if the social distance from $v$ to $d$ is no more than the one from $u$ to $d$.

We always assume that pairs $(u \rightarrow d)$ are allowed for all $u$, as any opportunity to complete the path with a single hop should not be missed. We have considered other rules as well (non-decreasing-degree, or strictly-decreasing-distance), which will be shown in a future version of this work.

Each rule above defines a heuristic method to select among all the opportunistic contacts the ones that are crucial to connect source and destination quickly over time. Our objective is to design a rule that minimize the contacts used, while allowing paths that approach the optimal delay.

For comparison we introduce two measures for each rule: its *selectivity*, which is measured by the fraction of the directed pairs $(u \rightarrow v)$ that it allows, and its *success ratio*, which is the probability that a path exists and follows this rule, with a maximum delay (it was set to ten minutes).

Typically this success can be normalized by the success of flooding.

The results for a selected set of rules may be found in Figures 6-7. We illustrate the delay distribution for two rules, as well as a comparison of selectivity and normalized success ratio (measured at ten minutes). For comparison, we have presented in some of them the performance (delay or success ratio) obtained when selecting contacts randomly according to the same selectivity.

The success ratio within ten minutes for flooding using all contacts is around 15%. When half of the contacts are selected randomly, the succcess ratio decreases to 10.5% (roughly 70% of flooding). In contrast, we observe, when contacts are selected based on social information, that one may either obtain a success ratio closer to flooding, with half of the contacts, or similar success ratio (around 70%) while selecting significantly fewer contacts. We have tested all rules mentioned above, and also combined several rules together (defined by intersection) to study their complementarity. The most important observations are as follows.

- The neighbor rule performs as well as most other rules (like, *e.g.*, non-increasing-distance); it performs significantly better than a random choice. This result is encouraging as people that are neighbors (*e.g.* friends, or friends of friends) are more likely to cooperate and construct a path over time.

- The rule based on centrality outperforms all the rules we have tested (reaching more than 95% of success with half of the pairs). It is better than the rule that uses the social distance to the destination, for similar selectivity. This result is counter-intuitive because this latter scheme is based on distance and depends on the destination, whereas non-decr.-centrality is destination unaware. Note that even the simpler rule based on degree outperforms the one based on distance.

- The combination of neighbor and centrality rules naturally improves selectivity, offering more flexibility and achieves one of the best trade-offs.

Our results indicate that, in the context of a community event like a conference, the primary factor to decide whether a node is a good next hop is its centrality. In addition, it
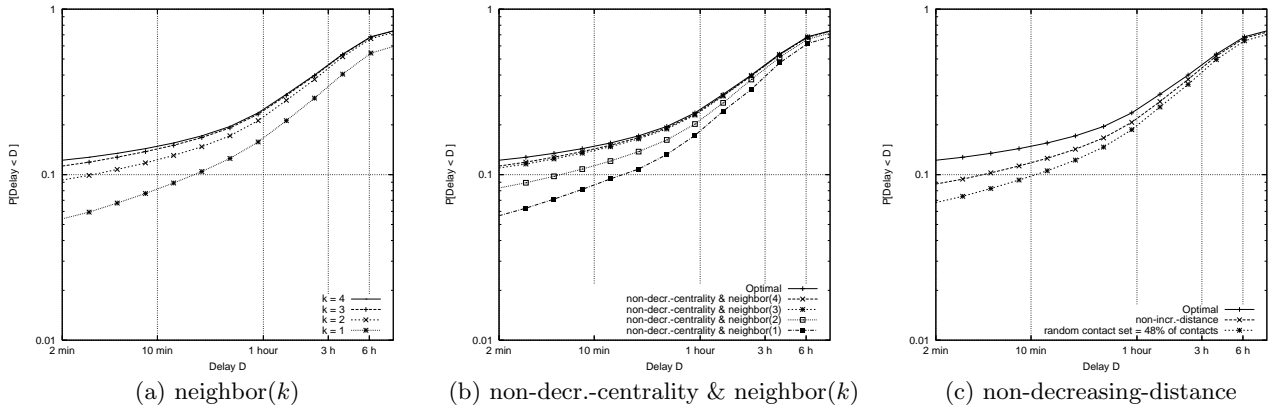
(a) neighbor(k)       (b) non-decr.-centrality & neighbor(k)       (c) non-decreasing-distance

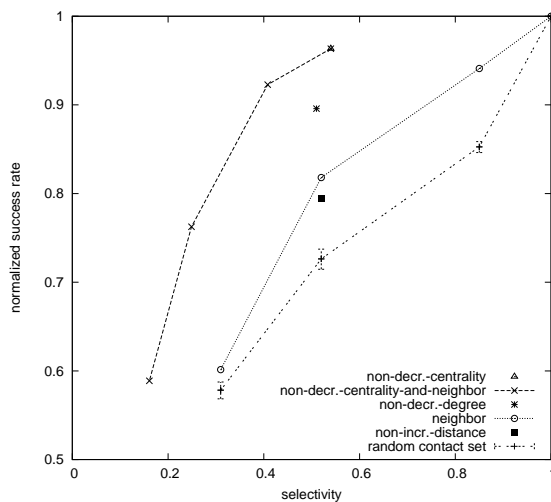**Figure 6: Performance of different path construction rule.**



**Figure 7: Comparison between path construction rules.**

seems that the best performance trade-off is obtained when several complementary rules are combined.

These encouraging results raise some important issues. First, in the absence of infrastructure, social information needs to be gathered locally by the social network application, and exchanged among the nodes. Some information (*e.g.*, node degree) can easily be computed locally, while other (*e.g.*, node centrality) depends on the global topology of the social network. In future work, we wish to see how to approximate the centrality of a node in a distributed manner, based on the social distance between the nodes. Second, path construction rules based on centrality can create congestion as they target always the same set of nodes. We hope that it could be mitigated by applying a complementary rule.

## 6. DISCUSSION

This paper presents initial comparison results of a social network, as defined by users in an OSN application, and measurement of opportunistic contacts between these users.

Our results, although limited to a single event happening inside a community, highlight the importance of central nodes. They also prove that using social neighbors to communicate (as in a traditional OSN) can be effective to exchange messages with opportunistic bandwidth.

In future work, we wish to extend the number of participants and the duration of the experiment to see how these results generalize to a larger group. This would also allows us to study how users belonging to different communities behave in a mobile social network application.

## 7. ACKNOWLEDGMENTS

## 8. REFERENCES

[1] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, J. Scott, and R. Gass. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Trans. Mob. Comp.*, 6(6):606–620, 2007.

[2] A. Chaintreau, A. Mtibaa, L. Massoulié, and C. Diot. The diameter of opportunistic mobile networks. In *Proc. of ACM CoNext*, 2007.

[3] E. M. Daly and M. Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *Proc. of ACM MobiHoc*, 2007.

[4] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: Social-based forwarding in delay tolerant networks. In *Proc. of ACM MobiHoc*, 2008.

[5] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic. Power law and exponential decay of intercontact times between mobile devices. In *Proc. of MobiCom*, 2007.

[6] D. Kempe, J. Kleinberg, and A. Kumar. Connectivity and inference problems for temporal networks. In *Proc. of ACM STOC*, 2000.

[7] J. Leskovec, J. Kleinberg, and C. Faloutsos. Graphs over time: densification laws, shrinking diameters and possible explanations. In *Proc. of ACM KDD*, 2005.

[8] A. Miklas, K. Gollu, K. Chan, S. Saroiu, P. Gummadi, and E. de Lara. Exploiting social interactions in mobile systems. In *Proc. of Ubicomp*, 2007.

[9] C. Tantipathananandh, T. Berger-Wolf, and D. Kempe. A framework for community identification in dynamic social nets. In *Proc. of ACM KDD*, 2007.