<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>The effect of communication pattern on opportunistic mobile networks</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Fan, X; Xu, K; Li, VOK; Yang, GH</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>The 8th IEEE Consumer Communications and Networking Conference (CCNC 2011), Las Vegas, NV., 9-12 January 2011. In Proceedings of the 8th CCNC, 2011, p. 1016-1020</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>2011</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/142817">http://hdl.handle.net/10722/142817</a></td>
</tr>
</tbody>
</table>

**Rights**

IEEE Consumer Communications and Networking Conference. Copyright © IEEE.; This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; ©2011 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
The effect of communication pattern on opportunistic mobile networks

Xiaoguang Fan, Kuang Xu, Victor O.K. Li, Guang-Hua Yang
Department of Electrical & Electronic Eng., The University of Hong Kong, Pokfulam Road, Hong Kong, China

Abstract—Social-based forwarding algorithms provide a new perspective on the study of routing in opportunistic mobile networks, and all of these schemes assume a uniform pattern for message generating rule. However, this is unconvincing due to the heterogeneity of contact rates in human communication patterns. In this paper we propose three social-based communication pattern models and utilize them to evaluate the network performance of different social-based routing protocols based on several human mobility traces. We find that communication patterns could significantly affect the network performance and the influence degree largely depends on the social metrics which these communication patterns are based on. We contend that considering communication pattern is quite important for designing a practical routing algorithm in opportunistic mobile networks.

I. INTRODUCTION

Human interactions have attracted much attention from sociologists, psychologists and anthropologists. It is non-trivial to make an accurate analysis on how and why certain interactions happen due to the uncertainty of human behavior and the complexity of communication patterns. Most sociologists agree that “the social environment in which an individual lives dramatically affects that individual’s behavior” [1]. Thus human interaction is far from fully determined by random factors. It is not a simple random communication, but consists of several building blocks, which interact with each other to influence human decisions on contact target and communication mode.

In an opportunistic mobile network, humans (nodes) use contact opportunities to realize information exchanging and message forwarding by short-range wireless connections without fixed network infrastructure. Each node needs to make forwarding decisions independently under store-carry-and-forward mechanism. Besides the epidemic routing and wait-for-destination scheme, many forwarding algorithms assign a “utility” metric to each node on making forwarding decisions. Messages are always forwarded to the node with higher utility. Moreover, social-based routing protocols like PROPHET [2], SimBet [3] and Fair Routing [4] utilize social characteristics of human interactions and implement social analysis technique on contact history to select the best forwarding target.

However, in simulation initialization, all these algorithms assume a uniform communication pattern on message generating rule, which means that each node in the network creates messages for any other nodes with the same probability. The assumption is unconvincing because in reality, a person may not contact others with the same frequency. By contrast, the heterogeneity of contact rates essentially changes the original information flow inside network and eventually determines the network status. Thus the simulation results under the assumption of uniform communication pattern may not reflect practical scenarios.

It is non-trivial to give an accurate description on human communication patterns. Due to the social characteristics of human interactions, the communication patterns should include some social context. In this paper, we intend to give a fundamental but methodical study on the effect of communication patterns on opportunistic mobile networks with various social-based forwarding algorithms. Here we summarize three social-based communication patterns: tie-strength-biased, centrality-biased and community-biased patterns. We firstly give a general introduction and model construction of these patterns, and then utilize different communication patterns and uniform pattern to simulate asynchronous messaging in real experimental human mobility traces. In the simulation we also implement three social-based routing protocols in order to evaluate both the general trend of network performance under different communication patterns, and the detailed variations between various forwarding algorithms.

We find that social-based communication patterns could further enhance the network throughput of different social-based routing protocols, and tie-strength-biased communication pattern offers the maximum performance enhancement. We also conclude that each forwarding algorithm does especially well under the communication pattern which is conceptually consistent with the heuristic of that forwarding algorithm. The implication is that we should utilize the social-based routing protocols selectively according to the communication patterns indicated in some specific scenarios.

We proceed in this paper as follows. Methodology is described in Section II, experiment design in Section III, simulation and evaluation in Section IV, and conclusion in Section V.

II. METHODOLOGY

Due to the complexity of human interactions in social environment, we try to analyze human communications from different aspects. Community, centrality and tie strength are three basic metrics in the theory of sociology. Community focuses on clusters of communications while centrality and tie strength tend to centralize attention on individual node and link within the network, respectively. In this paper, three kinds of communication patterns are proposed: community-biased, centrality-biased and tie-strength-biased, all of which are exhibited in real social environment.
A. Community-biased Pattern

Community is a fundamental concept in systematic sociology, since it is one of the basic elements in human society. Everyone in a society has a social position which illustrates his social identity and maintains his social characteristic. Community is a necessary and acceptable way to indicate one’s social role. Figure 1 shows the community distribution of one experiment at Cambridge University [5].

An organism of a given type might be more likely to interact with another organism of the same type than with a randomly chosen member of the population” [6], which can be formulated as the homophily principle [7], well studied in sociology [8]. We have done some elementary work on community-biased traffic in [9]. Here we further discuss its effect on routing in opportunistic mobile networks. We utilize the homophily principle to define community-biased communication pattern, following the intuition that an individual communicates with others in the same community more often than those in other communities. We use intra-community probability $P_{\text{intra}}$ and inter-community probability $P_{\text{inter}}$ to represent the possibility of a node generating messages to others in and out of his community, respectively. The sum of $P_{\text{intra}}$ and $P_{\text{inter}}$ should be 1.

B. Centrality-biased Pattern

Centrality is another way to demonstrate one’s social status as it reflects authority or popularity in a group [10]. Figure 2(a) shows the centrality distribution of nodes in the same experiment in Section II-A.

As an efficient way to measure one’s importance in a group, centrality plays a significant role in social-related analysis. If one is considered as a “star”, undoubtedly he is known by everyone and has a host of connections with others. As the centre of attention, a “star” always acts as an information hub bidirectionally: people tend to contact him and he would also like to associate with more persons to increase his importance. Therefore, individuals with high centrality value may occupy the majority of the communication traffic, which leads to the imbalance of contact rates.

In current literature, there are three widely used definitions for centrality with different angles on the topology of non-directional social graph, which are degree centrality, closeness centrality and betweenness centrality.

- **Degree centrality** focuses on the degree of an individual node, which can also be regarded as the total number of friends an individual has in the network. A node with high degree centrality has a great number of direct ties with other nodes and always acts as a cluster head in message broadcasting, forwarding and collecting [11].
- **Closeness centrality** is based on social distance. The major concern is how long it will take for a node to spread information to others in a social graph. Thus the utility metric of an individual node should be the reciprocal of total social distance from it to other nodes. [11]
- **Betweenness centrality** emphasizes the relay capability, or “interpersonal influence” [11], of nodes. The node with high betweenness value acts as an intermediate station to control the information traffic between nonadjacent nodes.

In our communication model, a node with high centrality value has a high probability to generate messages for others. The probability of node $n_i$ is calculated as:

$$P(n_i) = \frac{C(n_i)}{\sum_{k=1}^{N} C(n_k)},$$

where $C(n_i)$ is the centrality value of node $n_i$ and $N$ is the total number of nodes in the network.

Our major concern is how centrality influences the human communication mode. Degree centrality presses close to the real situation, e.g., the person who has many friends tends to make connections with more people. Thus we choose degree centrality as the metric in centrality-biased communication pattern.

C. Tie-Strength-biased Pattern

Tie strength tends to evaluate the social graph at a microscopic level. It concentrates on the robustness of relationship for a dyad [11], which represents a pair of nodes and the edge linking them. Large numbers of micro-level interactions could aggregate to form macro-level patterns which may feed back to individual dyads [13]. Thus tie-strength-biased communication pattern which is based on numerous interactions could greatly influence the total social graph. From the perspective of a node,

1The dataset we select includes 36 students equipped with bluetooth mobile devices for 11 days in Cambridge University.

2Betweenness of a node is the proportion of shortest paths between all possible pairs of nodes that pass through this node [12].
tie strength values are varying with different communication targets. Figure 2(b) shows the tie strength distribution of the same experiment in Section II-A.

Tie strength could be evaluated in various aspects, including frequency, recency and duration. Frequency focuses on the rate of recurrence for a number of contacts, while recency and duration concern the importance of one single contact. On the one hand, a single contact could be regarded as an important connection if it happens recently or lasts for a long period of time. On the other hand, a link between two nodes is believed to be strong if contact happens frequently on this link.

We illustrate our tie-strength-biased communication pattern following the intuition that a node would, with high probability, generate messages for the node who has strong tie with it. From the perspective of node $n_i$, the probability of sending messages to $n_j$ is calculated as:

$$P_{n_i}(n_j) = \frac{T(n_i,n_j)}{\sum_{k=1, k \neq i}^{N} T(n_i,n_k)},$$

where $T(n_i,n_j)$ is the tie strength value between nodes $n_i$ and $n_j$, and $N$ is the total number of nodes in the network. We choose frequency, the number of encounters a node has in a fixed period of time, as the measurement of tie strength.

### III. EXPERIMENT DESIGN

In this section, we illustrate the basic setting of our experiment, including two human mobility datasets and three typical social-based routing protocols in delay-tolerant network.

#### A. Dataset Description

We utilize two real human mobility traces collected by two research projects, Reality Mining [14] at MIT and Hagggle [5] at Infocom2006 conference. We denote them as MIT reality mining and Infocom2006 respectively. In these experiments, Bluetooth-enabled mobile devices logged contacts with each other by doing Bluetooth device discovery periodically. In MIT reality mining, the data session we use is from March 1 to March 25 in 2005 as it does not include long term holidays. In Infocom2006, we choose a five-hour period from 8:00 am to 1:00 pm on April 24 2006, which is a busy period including meeting and lunch time.

The post-process datasets we use in this paper are summarized in Table I. Here we define contact density and contact diversity to further compare the properties of two datasets. Obviously, contact density represents the contact frequency of the whole group, which is the average number of daily contacts per pair. Considering the expression of dyad in Section II-C, we define contact diversity as the fraction of connected dyads out of all possible pairs in the network, including both connected dyads and the pairs that have no links. From table I, we can see that MIT reality mining is very sparse with low contact density and diversity, while Infocom2006 is an intensive opportunistic network due to its high level of contact aggregation in meeting mode. Figure 3 shows the contact diversity distribution of these two datasets.

We can see that most of nodes never encounter 50% of the total population in MIT reality mining. Later we shall see how these properties affect network performance under various communication patterns.

<table>
<thead>
<tr>
<th>Experimental dataset</th>
<th>MIT Reality</th>
<th>Infocom2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of devices</td>
<td>80</td>
<td>78</td>
</tr>
<tr>
<td>No. of contacts</td>
<td>4000</td>
<td>30000</td>
</tr>
<tr>
<td>Average No. of contacts/pair</td>
<td>1.266</td>
<td>9.990</td>
</tr>
<tr>
<td>Contact density</td>
<td>0.0452</td>
<td>47.952</td>
</tr>
<tr>
<td>Contact diversity</td>
<td>17.5949%</td>
<td>57.6091%</td>
</tr>
</tbody>
</table>

Table I: FEATURES OF TWO HUMAN MOBILITY TRACES

![Figure 3. Contact Diversity for MIT Reality Mining and Infocom2006.](image)

#### B. Forwarding Algorithm

Three distributed social-based forwarding strategies are evaluated in our simulation, all of which operate in individual node and intend to find the best next hop by comparing the utility metrics with different encounter nodes.

**PROPHET** [2] utilizes contact frequency to define a probabilistic utility called delivery predictability which increases a certain amount by instant encounter and decreases exponentially with time. The calculation of delivery predictability also concerns the transitive property derived from the notion of weak tie [13]. In other words, node $A$ tends to forward message to node $B$ if there is a common friend between them, who contacts both of them frequently. Thus PROPHET is based on the idea that social links with high strength value may be a good choice for message forwarding. We set the parameters $P_{init} = 0.75$, $\beta = 0.25$, $\sigma = 0.98$ according to the author’s suggestion.

**SimBet** [3] associates betweenness centrality with similarity in the evaluation of utility. Here similarity is defined as the total number of common neighbours between individuals. A node is highly favored to be the forwarding target if it has large betweenness centrality and similarity value. The goal is to use similarity to identify community and betweenness centrality to find bridges between these communities. Messages are first transferred between communities through central nodes and then delivered to the target within the destination community. We set the parameter $\alpha = 0.5$ according to the author’s advice.

**FairRouting** [4] provides two estimators, short term and long term robustness, to indicate perceived interaction strength at different time scales. When defining utility, it uses the aggregated interaction strength, which is proportional to both the long term robustness and the difference between long term and short term robustness value, to identify sustainable long
term tie by picking out ephemeral relationship. Moreover, it controls the queue size of each node to balance the traffic load.

IV. SIMULATION AND EVALUATION

A. Simulation Setting

We build a contact-driven simulation platform to evaluate routing with mobility traces. The metric of network performance we use is system throughput, which is defined as the proportion of successfully delivered messages out of all generated messages. In the message generating process, each node will select a list of destinations according to certain communication pattern. Due to the lack of priori community information, a community detection tool CFinder [15] is applied to each traceset. Here we intend to divide nodes into two communities. Since not all nodes can be subsumed in community and some nodes may lie in both of two communities, we also assume that the node which belongs to neither of two communities generates messages in a uniform pattern, and the node which lies in the overlapped part generates messages to the members of both communities with intra-community probability.

Considering the contact aggregation problem [16] as well, we set a threshold to identify the linking. i.e., nodes are supposed to be friends only when the number of contacts between them exceeds a threshold. The reference threshold value is the average number of contacts per pair shown in Table I.

B. Result and Discussion

We study the delivery success rate from different angles to show the effect of communication patterns on opportunistic mobile networks and evaluate different social-based forwarding algorithms under the context of communication patterns.

Figure 4 shows the system throughput of three social-based routing protocols under different communication patterns in the MIT reality mining dataset. The x-axis represents the number of contacts. One observation is that the order of delivery success rates for three social-based forwarding algorithms are different under different communication patterns. Figure 4 (a) represents the uniform pattern. PROPHET is the best at first, but is surpassed by SimBet after 3300 contacts. However, under community-biased and centrality-biased patterns, shown in Figure 4 (b)(c), SimBet is better than PROPHET and FairRouting. As stated in Section III-B, the core idea of SimBet is based on centrality and community. Thus the centrality-biased and community-biased communication patterns would significantly enhance the performance of SimBet due to the conceptual consistency of the social metrics they utilize. Furthermore, in tie-strength-biased pattern shown in Figure 4 (d), PROPHET outperforms SimBet. The throughput of FairRouting is also close to SimBet, which indicates that tie-strength-biased communication pattern greatly improves the performance of PROPHET and FairRouting, since they are both set up on the strength evaluation of social ties. We conclude that social-based communication patterns could significantly influence the network performance of different social-based routing protocols and each forwarding algorithm does especially well under the communication pattern which is conceptually consistent with the heuristic of that forwarding algorithm.

Figure 5 shows the comparison of delivery success rates under four communication patterns. We choose SimBet and PROPHET as two selected forwarding algorithms and show their performance under the MIT reality mining and Infocom2006 datasets, respectively. We observe that all the social-based communication patterns enhance the network performance when compared with the uniform pattern. This is because the original network characteristics significantly coincide with the social features of these communication patterns. We also find that different communication patterns affect the system throughput in a fixed order. Under both SimBet and PROPHET algorithms, which are shown in Figure 5 (a)(c) and (b)(d), respectively, tie-strength-biased communication pattern offers the best performance. This indicates that tie strength should be the major concern when considering routing in opportunistic mobile networks. Another interesting observation is that community-biased pattern is better than centrality-biased pattern in Figure 5 (a)(b), but the order is reversed in Figure 5 (c)(d). This is attributable to the specific characteristics of the two real human mobility traces. According to the dataset analysis in Section III-A, MIT reality mining is very sparse and short of contact diversity, while Infocom2006 has higher contact frequency and the contacts cover the majority of node pairs in the network. Thus we conclude that the network topology, which is indicated by the contact density and diversity, could influence the network performance significantly under centrality-biased and community-biased communication patterns. In the social graph with sparse contact frequency and low contact diversity, the network throughput under community-biased communication pattern is higher than centrality-biased communication pattern. By contrast, the social graph with dense contact frequency and high contact diversity makes centrality-biased communication pattern work better.

Finally we summarize the most general observations from the study of the effect of communication patterns on routing in opportunistic mobile networks.

- Social-based communication patterns significantly increase the system throughput of social-based routing protocols when compared with uniform communication pattern, and tie-strength-biased communication pattern offers the best performance.
- Each social-based forwarding algorithm could further enhance the performance under the communication pattern which is conceptually consistent with the heuristic of that forwarding algorithm.
- The network topology, as indicated by the contact density and diversity, could greatly influence the network performance under centrality-biased and community-biased communication patterns.

3Fair Routing has the same trend as the other two algorithms and we left it out due to space limitations.
Thus the implication is that we could utilize social-based routing protocols selectively according to the communication patterns indicated in some specific scenarios. Some scenarios reflect concentration on group talking and there are several central individuals who occupy the majority of communication traffic. For example, in conference or workshop, people aggregate to discuss for a long period of time and famous researchers are always the focused persons. In this situation, a routing protocol like SimBet may be utilized due to its compatibility with community-biased and centrality-biased communication patterns. By contrast, some applications depend critically on the quality of connections. For example, in social recommendation system, people only want to share the material with their best friends. In this situation, PROPHET may be a good choice since tie-strength-biased communication pattern favors more on tie-strength-based forwarding algorithms.

V. CONCLUSION AND FUTURE WORK

This paper is a preliminary study on the effect of communication patterns on opportunistic mobile networks. We believe that our observations could be widely applied in the assessment of forwarding algorithms, and we contend that considering communication pattern is quite important for designing a practical routing algorithm in opportunistic mobile networks. Based on the empirical analysis, we will further build an evaluation platform under the assumption of communication patterns in order to assess the performance of different routing protocols and select compatible forwarding schemes in some specific practical scenarios.

REFERENCES