

Geocommunity-Based Broadcasting for Data Dissemination in Mobile Social Networks

Jialu Fan, *Member, IEEE*, Jiming Chen, *Senior Member, IEEE*, Yuan Du, Wei Gao, *Student Member, IEEE*, Jie Wu, *Fellow, IEEE*, and Youxian Sun

Abstract—In this paper, we consider the issue of data broadcasting in mobile social networks (MSNs). The objective is to broadcast data from a superuser to other users in the network. There are two main challenges under this paradigm, namely 1) how to represent and characterize user mobility in realistic MSNs; 2) given the knowledge of regular users' movements, how to design an efficient superuser route to broadcast data actively. We first explore several realistic data sets to reveal both geographic and social regularities of human mobility, and further propose the concepts of *geocommunity* and *geocentrality* into MSN analysis. Then, we employ a semi-Markov process to model user mobility based on the geocommunity structure of the network. Correspondingly, the geocentrality indicating the “dynamic user density” of each geocommunity can be derived from the semi-Markov model. Finally, considering the geocentrality information, we provide different route algorithms to cater to the superuser that wants to either minimize total duration or maximize dissemination ratio. To the best of our knowledge, this work is the first to study data broadcasting in a realistic MSN setting. Extensive trace-driven simulations show that our approach consistently outperforms other existing superuser route design algorithms in terms of dissemination ratio and energy efficiency.

Index Terms—Mobile social networks, data dissemination, broadcasting, geography, community

1 INTRODUCTION

MOBILE Social Networks (MSNs) are networks in which mobile social users physically interact with each other and further reach network service, even in the absence of network infrastructure or end-to-end connectivity [1], [2]. MSNs can be viewed as a kind of socially-aware Delay/Disruption Tolerant Networks (DTNs). Thanks to the popularization of smartphones (e.g., iPhone, Nokia N95, and Blackberry), MSNs have begun to attract more attention. However, intermittent and uncertain network connectivity make data dissemination in MSNs a challenging problem.

Broadcasting is the operation of sending data from a source user to all other users in the network. This is frequently used in many applications of mobile ad hoc networks (MANETs) [3], [4]. On the other hand, the existing work in intermittently connected networks always focuses on data unicast [5] or multicast [6]. However, broadcast is more effective for data dissemination in such opportunistic environments. Most of the envisioned services (ranging

from safety applications to traffic management [7]) rely on broadcasting data to the users inside a certain area of interest. For example, location-based services (product prices, tourist points of interest, etc.) can be advertised from salesmen to near-by users.

In this paper, we focus on data broadcasting from a single, special user to all other users in MSNs. Specifically, the special mobile user is called *superuser*, and the other regular users are called *users* for short. With the knowledge of users' movements, data broadcasting depends mainly on the mobility trajectory of the superuser. Therefore, the design of a superuser route has a significant impact on network performance. Although there is some similar work on special user route design, such as Message Ferry [8], [9] or Data MULEs [10], they always assume that the special users move with fixed routes [8] to facilitate connectivity among other users, or aim to deal with energy conservation rather than data transmission [10]. The ferry trajectory in SCFR [11] is semideterministic, depending on the traffic rate. However, it considers the network with static nodes. Tariq et al. [12] proposed a customized ferry route for mobile networks, but the node mobility is strongly constrained. Besides, none of them considers the characteristics of user mobility in realistic MSNs.

Our primary goal is to design flexible superuser routes for data broadcasting in MSNs, without any constraints on the movements of regular users. Hence, the main challenge is how to characterize and represent user mobility in MSNs. From a social network perspective, people sharing interesting properties (e.g., common hobbies, social functions, and occupations) tend to form a *community*. Through trace-based study, we detect an interesting phenomenon: *in MSNs, community always strongly relates to geography*. For example, graduate students working in the

• J. Fan is with the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, P.R. China. E-mail: fanjialu@gmail.com.

• J. Chen, Y. Du, and Y. Sun are with the State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, P.R. China. E-mail: {jmchen, yxsun}@iipc.zju.edu.cn, duyuan@zju.edu.cn.

• W. Gao is with the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville, TN 37996. E-mail: weigao@utk.edu.

• J. Wu is with the Department of Computer and Information Sciences, Temple University, 1805 N. Broad ST., Room 302, Wachman Hall 302, Philadelphia, PA 19122. E-mail: jiewu@temple.edu.

Manuscript received 12 May 2011; revised 25 May 2012; accepted 27 May 2012; published online 6 June 2012.

Recommended for acceptance by S. Papavassiliou.

For information on obtaining reprints of this article, please send e-mail to: tpds@computer.org, and reference IEEECS Log Number TPDS-2011-05-0302. Digital Object Identifier no. 10.1109/TPDS.2012.171.

same office form a community, and they always contact each other in the office. Therefore, we propose *geocommunity*, which represents a geography-related community, with MSNets as a fundamental structure. By means of geocommunity, we characterize user mobility and further design superuser route to actively broadcast data to mobile social users in the network.

The novelty and contributions of this paper are as follows:

- We use three data sets collected from realistic MSNet environments to study the characteristics of user mobility. The experiment results show that people in a human society also express geographic regularity, as a supplementary social attribute. Therefore, we propose *geocommunities* into MSNets to characterize both geographic and social regularities of user mobility.
- Through trace-based study, we detect that a user's sojourn time at a geocommunity does not follow the exponential, but instead a power-law distribution. Hence, we formulate user mobility over geocommunities in MSNets as a semi-Markov model.
- With a semi-Markov model, we compute each user's steady-state probability distribution over geocommunities, and further propose *geocentrality* to measure the user density of each geocommunity.
- Considering geocentrality, we propose *Static Route Algorithms (SRA)* from a statistic perspective to the superuser that wants to either minimize total duration of the route (*min-T-SRA*), or maximize dissemination ratio (*max-p-SRA*). Furthermore, we also propose a *Greedy Adaptive Route Algorithm (GARA)* excluding the overlap of contact user sets (CUS) among the geocommunities.

The remainder of this paper is organized as follows: Section 2 provides an overview of two trace-based observations in MSNets, as well as the big picture and the optimization objectives of our data broadcast scheme. Then, we explore both social and geographic regularities of user mobility in realistic traces, and further propose the concepts of geocommunity and geocentrality into MSNet analysis in Section 3. Based on two such properties, we employ a semi-Markov process to model user mobility in Section 4, and propose superuser route schemes in Section 5. Section 6 evaluates the performance of our approach, via realistic trace-driven simulation. The last two sections present related work and conclusions, respectively.

2 OVERVIEW

In this section, we first introduce two trace-based observations from practical MSNet environments to motivate our proposed broadcast scheme. Then, we give a big picture of the scheme, i.e., the superuser route design problem, whose optimization objectives are finally presented.

2.1 Two Observations from Realistic MSNet Traces

Through trace-based study, we have the following two observations:

- *Mobile users in MSNets usually move around several well-visited locations instead of moving randomly.* We

explore the realistic traces aiming to reveal the real situation behind such observation. The well-known "small world" phenomenon shows that people usually belong to several communities, and contact others with similar hobbies, occupations, or social functions. For example, graduate students working in the same office interact more frequently with each other. Based on the community concept, we further detect that such contact preference is also usually correlated to geography information, such that the contacts among officemates mostly happen in the office. We define such a geography-related community as a *geocommunity*, which will be experimentally explored in Section 3.3. Furthermore, the user's sojourn time at each geocommunity is fairly regular, because their social behavior patterns usually remain stable within a relatively long interval [13].

- *Spatial user distributions are very heterogeneous and possess several geocommunities of high user density.* Since geocommunity affiliations among mobile users can be highly diverse, MSNets have some geocommunities of higher user density, and where the superuser therefore has a much better chance of encountering regular users than elsewhere. Examples of such geocommunities include public transportation and shopping centers in urban environments, or conference rooms and cafeterias in office buildings, etc. Therefore, we also propose *geocentrality*, a geography-related centrality metric, into MSNets to measure the user density of geocommunities. Such a metric will be described in detail in Section 3.4, and will be further used in the superuser route design.

2.2 Big Picture

We consider a following scenario: a salesman is about to advertise his products to on-campus customers (e.g., faculty, staffs, and students). He has to physically move around the campus to transmit advertisements via his smartphone to users'. He is trying to decide his route, aiming to broadcast the ads to mobile users as soon as possible.

In our data broadcast scheme, the users in MSNets are classified into two categories: 1) regular users, or simply, the users that move based on their social lives. These users are potential data receivers from the superuser. The movements of the users are not controllable; 2) a single, special user called superuser (e.g., the salesman) that aims to broadcast data to regular users in the network. In this paper, we only consider *one-hop* data broadcasting from the superuser to regular users. Opportunistic transmission between regular users is not our focus. This is because rational users can be expected to behave selfishly, and need incentive to cooperate in human society; because of this, we believe the incentive schemes in MSNets deserve separate studies. The data broadcast problem is, therefore, how to design a superuser route to facilitate data broadcasting most effectively. To solve such a problem, we should focus on answering the following questions:

- What is the appropriate metric for measuring the dynamic user density of geocommunities?

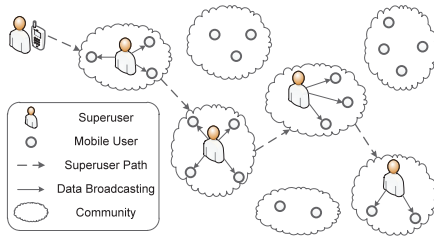


Fig. 1. Diagram demonstrating our geocommunity-based data broadcast scheme in MSNets.

- Given the dynamic user density of each geocommunity, how should the superuser decide which geocommunities can stay, and for how long, respectively?

Generally speaking, the exploitation of both social and geographic regularities of user mobility in MSNets will definitely facilitate the calculation of geocentrality metric. As a matter of fact, users in MSNets always move around several well-visited geocommunities, and their sojourn times at each geocommunity remain stable over time. Hence, we employ a Markov process to model user mobility in the network, where the geocommunities are represented as Markovian states.

Through analyzing the Markov model, we can compute each user's steady-state probability distribution over geocommunities, and further propose geocentrality to measure each geocommunity's dynamic user density. Suppose the whole network is composed of a certain number of geocommunities. The problem of superuser route design is then turned to: How should the superuser choose geocommunities, and allocate waiting times accordingly? Fig. 1 gives an illustration of our geocommunity-based data broadcast scheme.

2.3 Optimization Objectives

We consider two cases: 1) the superuser aims to minimize the total duration of route T under the constraint of a certain dissemination ratio, which corresponds to *time-sensitive superuser*. For example, the salesman tries to use less time to advertise the product to a certain number of people; 2) the superuser aims to maximize the dissemination ratio p before a certain deadline, which corresponds to *dissemination-ratio-sensitive superuser*. In this case, the salesman tries to advertise the product to more people within working time.

In both cases, the superuser route design scheme follows a utility-based approach. The superuser route comprises some geocommunities and the corresponding waiting times. Suppose that the whole network is composed of a certain number of geocommunities, and the superuser can immediately transmit data to all the users within the geocommunity where the superuser is stopping. The utility u_i of geocommunity i describes its potential contribution to the superuser's data dissemination. From a statistic perspective, the number of users to whom the superuser transmits data at geocommunity i does not decrease with increasing t_i . This indicates the superuser's waiting time at geocommunity i . In other words, $u_i(t_i)$ is a nondecreasing function of t_i . However, the key point is that the increment of $u_i(t_i)$ with t_i is different among geocommunities;

TABLE 1
Trace Summary

Trace	<i>MIT Reality</i>	<i>Infocom 06</i>	<i>CoSphere</i>
Device	Smart Phones	iMote	Smart Phones
Network type	Bluetooth	Bluetooth	Cellular, 802.11 and Bluetooth
Duration (days)	246	3	42
Granularity (minutes)	5	2	5
No. of devices	97	78	12

therefore, we should then allocate the limited time to geocommunities with higher gradient of $u_i(t_i)$. It will be shown later that the associated utility of a geocommunity is strongly related to its geocentrality.

3 TRACE-BASED ANALYSIS ON MOBILE SOCIAL NETWORKS

In this section, we explore both social and geographic regularities of user mobility in MSNets based on realistic trace study, and further propose the concepts of geocommunity and geocentrality.

3.1 Experimental Traces

We study the characteristics of user mobility on three sets of MSNet traces. We believe that the chosen traces cover a rich diversity of environments, from crowded conference sites (*Infocom '06*) [14] to quiet university campuses (*MIT Reality* and *CoSphere*) [15], [16], with experimental periods from a few days (*Infocom '06*) to almost one year (*MIT Reality*). The three traces are summarized in Table 1.

We choose traces containing static Access Points (APs) because APs have geographic properties. Specifically, we use the *syslog* data for mobile users' association patterns to APs. From these *syslog* messages, the mobility of each user is extracted in the form of a series of two tuples (AP name and the time stamp when the association occurs). For simplicity, we classify more than 30,000 APs in the *MIT Reality* trace into 50 geolocations. In both *Infocom '06* and *CoSphere*, the neighborhood of each AP corresponds to one geolocation.

3.2 Geographic Regularity of User Mobility

We investigate whether or not, and to what extent, user mobility behaviors in realistic MSNets correlate in time and space. From a user-centric view, we define the "contact geolocation set (CLS)" as follows:

Definition 1. The *Contact geolocation Set (CLS)* of a user i during time period $[t_1, t_2]$ is a geolocation set \mathcal{L}_i , where for any $1 \leq j \leq |\mathcal{L}_i|$, the cumulative sojourn time $T_{i,j}$ of user i at geolocation j is larger than λ , which is a predefined threshold.

Intuitively, the CLS of a user is comprised of several geolocations where he/she always visits. Note that the introduction of λ is in order to exclude the effect of pass-by geolocations to some extent. First, we compare the similarity of user's sojourn time distributions over CLS, based on traces collected in time intervals of different scales from the *MIT Reality* data set, where λ is set to 0. We choose *cosine distance* as the similarity measure, since it is always used to measure similarity in the field of text and data

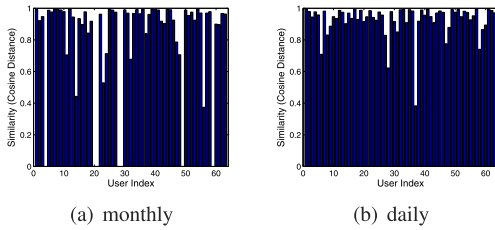


Fig. 2. Similarity between users' *CLS*, derived based on the *MIT Reality* traces of different timescales.

TABLE 2
Numerical Parameters on the *CLS* Size Distribution

Trace	<i>MIT Reality</i>	<i>Infocom 06</i>	<i>CoSphere</i>
Mean (μ)	3	1.7846	2.1818
Variance (σ^2)	1.0389	1.2464	0.9360

mining [17]. In the current problem, the cosine distance $sim(\vec{p}, \vec{q})$ is defined as

$$sim(\vec{p}, \vec{q}) = \frac{\sum_{j=1}^J p_j q_j}{\left(\sqrt{\sum_{j=1}^J p_j^2}\right) \left(\sqrt{\sum_{j=1}^J q_j^2}\right)} = \frac{\vec{p} \cdot \vec{q}}{|\vec{p}| |\vec{q}|}, \quad (1)$$

where p_j and q_j indicate a user's sojourn time at geolocation j during the two time periods under comparison. Then, $\vec{p} = [p_j]$ and $\vec{q} = [q_j]$ are the sojourn time distributions over the user's *CLS* during the two corresponding time periods, respectively. Note that $sim(\vec{p}, \vec{q})$ ranges in $[0, 1]$, with $sim(\vec{p}, \vec{q}) = 1$, indicating that the user's mobility behaviors in the two time periods are identical.

Fig. 2a shows the cosine distance between monthly traces collected in a period of two months (October and November, 2004), where the gaps mean the corresponding users have no *syslog* in this period. It can be observed that over 85 percent of active users¹ have similarity values higher than 0.9, most of which approach 1. Fig. 2b gives such a similarity measure, derived based on daily traces. As the mobility data collected from the daily trace is insufficient to characterize the mobility behavior, we construct the mobility models at the daily scale by separately gathering all the *syslog* data on Monday and Tuesday during the whole nine-month experiment period. It is again observed that the similarity values are extremely high in most users. In short, the results show that users in MSNets always keep a stable movement schedule in a relatively long interval. Hence, we can extract a user's past mobility characteristics to predict his/her future movement.

We also study the *CLS* size of each user, which represents the number of geolocations that a user mostly moves around during the experiment period. Here, λ is set as 30 min per day to exclude the effect of pass-by geolocations. The results in Table 2 show that the mean value of users' *CLS* size is 3 in the *MIT Reality* trace. On the other hand, such measures in *Infocom '06* and *CoSphere* are even lower, around 2. The low variances mean that the *CLS* size of all the users stays around the mean. Hence, it can be concluded that users in MSNets usually move around

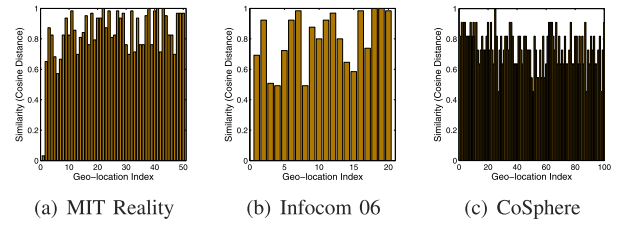


Fig. 3. Average similarity of *CUS* in different timescales for each geolocation.

several locations instead of all over the whole network. It also demonstrates the phenomenon of skew spatial distributions of users in MSNets.

3.3 Geocommunity

A community is defined as a clustering of users that are "tightly" linked to each other, either by direct linkage or by some "easily accessible" users that act as intermediates. Members of a community usually share interesting properties, such as common hobbies, social functions, and occupations. On campus, graduate students working in the same office interact more frequently with each other; members affiliating with the same team, such as football or swimming, have such heavy interactions.

Intuitively, interests might relate to geography in human society²: officemates contact each other in the office; volleyball lovers play volleyball together in gyms. It follows that we experimentally study the correlation between community and geography in the three data sets, in terms of the stability of "contact user set" of each geolocation, and spatial distributions of pairwise user contacts. Similar to *CLS*, we define the following "contact user set" from a location-centric view:

Definition 2. The *Contact User Set* of a geolocation j during time period $[t_1, t_2]$ is a user set \mathcal{S}_j , where for any $1 \leq i \leq |\mathcal{S}_j|$, the cumulative sojourn time $T_{i,j}$ of user i at geolocation j is larger than λ , which is a predefined threshold.

First, we explore the stability of the *CUS* of a geolocation during the whole experiment period. Here, λ is still set to 30 min per day. We use cosine distance to measure the similarity of *CUS* at the same geolocation between different time periods. As seen from Fig. 3, the mean value of such similarity measures is 0.8 in the *MIT Reality* trace on a monthly basis, and are both around 0.75 on a daily scale in the *Infocom '06* and *CoSphere* traces, respectively. The high similarity values prove that the contacting member structure of a geolocation in MSNets keeps stable in an extremely long period.

Then, we also study the spatial distributions of pairwise user contacts. We choose one user with the largest size of *syslog* from the traces to analyze the distribution of geolocations where he/she makes contact with the other users in the network. Fig. 4 gives the standard deviation of contact counts at different geolocations. As seen from Fig. 4a, such a measure is extremely high in the *MIT Reality* trace, and over 79 percent of other users contact the chosen

1. Every user visits at least one geolocation during the experiment duration.

2. Online social networks, such as Facebook and MySpace, are beyond the scope of this paper.

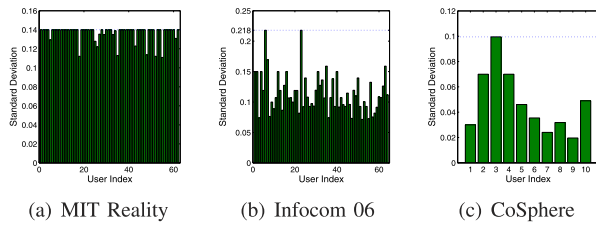


Fig. 4. Standard deviation of spatial distribution of pairwise user contacts over geolocations.

user at only one geolocation, i.e., the standard deviation reaches the upper bound accordingly. The real situation is that a user in MSNets affiliates to several different communities, and then they periodically act as the different roles, respectively. Hence, a user usually contacts other users who have at least one common interest with him/her at the according geolocations.

Therefore, we propose a geography-related community, *geocommunity*,³ to characterize the stable-member structure of a geolocation in MSNets. For simplicity, we assume that one geolocation corresponds to a geocommunity, and we consider the sojourn time of a participant spent at a geocommunity to be the time interval of his/her two consecutive contacts with different geolocations, the former of which is the corresponding geocommunity. Later, the geocommunity shows its efficiency in modeling user mobility in MSNets.

A single user can be a mutual member of several different communities (people have varying roles in society). However, we consider that a user only shows one membership at any single time in this paper. The user can change geocommunity affiliation over time, and we assume he/she does not spend any time on the transition between geocommunities.

3.4 Geocentrality

“Betweenness” centrality, which measures the extent to which a node lies on the paths linking other nodes, is currently a widely used Freeman’s measure in social-based data forwarding [5]. Certainly, betweenness is not sufficient to analytically represent the probabilities for a geocommunity to contact mobile users in the network. Inspired by [6], we propose geocentrality, a centrality metric of the geocommunity, into MSNets:

Definition 3. Suppose there are a total of N users in the network, and the steady-state contact probability per unit time between geocommunity i to $User_k$ is ϕ_i^k . The *geocentrality* of geocommunity i is defined as

$$C_i(t_i) = 1 - \frac{1}{N} \sum_{k=1}^N (1 - \phi_i^k)^{t_i}, \quad (2)$$

where $C_i(t_i)$ indicates the average probability that a randomly chosen user in the network is contacted by geocommunity i within time t_i . The unit time means we focus on the discrete time system in our work. A user being contacted by a geocommunity indicates that the user affiliates with that community.

3. We will use the terms, *community* and *geocommunity*, interchangeably in subsequent sections.

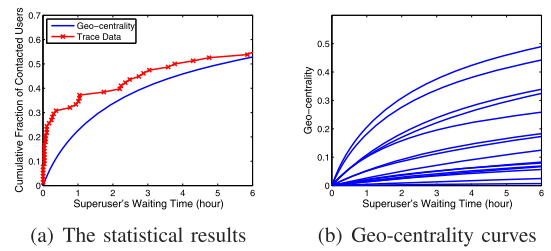


Fig. 5. Empirical study on geocentrality in *Infocom '06*.

To show the effectiveness of the geocentrality metric in characterizing the capability of a geocommunity to contact mobile users in the network, we choose a geocommunity in the *Infocom '06* trace and run it 100 times with a random starting time for statistical convergence. The statistical results on the number of contacted users show the trend to form the red line in Fig. 5a, where the geocentrality curve is drawn with the unit time equal to 3 minutes. The comparison results prove the feasibility of using geocentrality to represent the dynamic user density of a geocommunity. We further draw the geocentrality curves of all the geocommunities in the *Infocom '06* trace in Fig. 5b, where it can be observed that geocommunities have the heterogeneous capability to contact users, i.e., spatial user distributions over geocommunities are very diverse.

4 USER’S MOBILITY MODEL

Users in MSNets always belong to several communities; hence, they usually move around these well-visited locations (i.e., geocommunities). Therefore, we can model user mobility as a Markov process, where the stated space can be represented by the set of geocommunities. In this section, we first explore the sojourn time distributions of users over geocommunities in MSNets. The resulting power-law distribution enables a semi-Markov modeling; hence, we experimentally evaluate the feasibility of modeling user mobility as a semi-Markov process with realistic traces.

4.1 User’s Sojourn Time Distribution over Geocommunities

We first study the distributions of a user’s sojourn time over geocommunities on the three traces. As seen in Fig. 6, the sojourn time distributions approximate to the power law within a certain range (the approximate lines are drawn with red lines). Since a power law is determined by the gradient of the line on log-log graphs, the coefficient partly characterizes the sojourn time distribution of the traces. For *Infocom '06*, the coefficient is 1 in the range [5 min, 1 h], while the distribution coefficient of *CoSphere* is 0.95 in the range [10 min, 2.5 d].

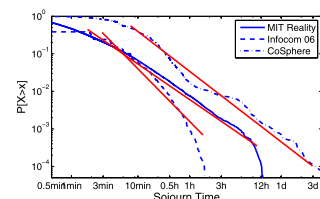


Fig. 6. Distributions of sojourn time for the realistic traces.

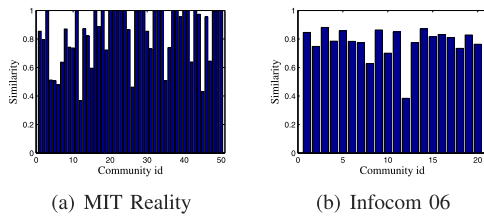


Fig. 7. Similarity between semi-Markov steady-state distributions and realistic statistical distributions.

It can also be observed that the sojourn time distributions have a heavy tail, i.e., the sojourn time decreases slowly in the tail. In the *Infocom '06* trace, around 15 percent of records are more than 4.5 minutes, and over 5 percent are larger than 12 minutes; for the *MIT Reality* trace, 20 percent of contacts last more than 3 minutes, and 5 percent are longer than 13 minutes.

Since a user's sojourn time over geocommunities follows power law, but not exponential distributions; we employ semi-Markov process instead of continuous-time Markov chain to model user mobility in MSNets. The specific computation methods of several critical parameters of a semi-Markov model are given in [18]. In this paper, we focus on the steady-state distribution ϕ_i^k of the model to compute geocentrality of communities, and further design algorithms described in Section 5 for the superuser route.

4.2 Similarity of Users' Steady-State Distributions in Different Timescales

We experimentally evaluate the feasibility of modeling user mobility in MSNets as a semi-Markov process. We, respectively, calculate the semi-Markov steady-state distributions ϕ_i^k and the average contact probability of users over geocommunities based on realistic traces. The average contact probability between user and his/her affiliated geocommunities can be statistically obtained from the realistic traces. We then calculate the cosine distance between the two metrics, with the aim to investigate to what extent the steady-state distributions derived from the semi-Markov model represent the actual situation.

As seen from Fig. 7, most similarity values are more than 70 percent in both traces. Specifically, some values almost reach 100 percent. The results verify that it is appropriate to model user mobility in MSNets as a time-homogeneous semi-Markov model, as far as the long-term movement behavior is concerned.

We also calculate the steady-state distributions of users over geocommunities, based on the *MIT Reality* traces collected in time intervals of different scales. The aim is to further investigate to what extent the steady-state distributions derived from the semi-Markov model correlate in time. We collect traces on monthly and daily bases to derive different steady-state distributions of users over geocommunities, and compare them with cosine distance as described in (1). Here, $\vec{p} = [p_j]$ and $\vec{q} = [q_j]$ are the steady-state distributions of two semi-Markov models under comparison, respectively. Table 3 gives the above similarity measure between steady-state distributions derived based on traces of *MIT Reality* in different timescales. The monthly traces are collected in a period of seven months (28-week

TABLE 3
Similarity between Steady-State Distributions $[\phi_i^k]$ Derived Based on *Monthly/Daily* Traces

Monthly	<i>m2</i>	<i>m3</i>	<i>m4</i>	<i>m5</i>	<i>m6</i>	<i>m7</i>
<i>m1</i> v.s.	0.9693	0.9405	0.3183	0.8366	0.9924	0.9270
Daily	<i>Mon</i>	<i>Tue</i>	<i>Wed</i>	<i>Thu</i>	<i>Fri</i>	<i>Sat</i>
<i>Sun</i> v.s.	0.7077	0.7347	0.6937	0.6700	0.2898	0.5709

period) between September 19, 2004 and March 4, 2005. It can be observed that the closer in time the two monthly traces, the more similar the corresponding two steady-state distributions; though, there exist some exceptions. For example, the cosine distance becomes much smaller when *m4* is involved. This is because *m4* corresponds to the period of *winter break* (from December 12, 2004 to January 8, 2005). This administrative event on campus affects user mobility dramatically. It is again confirmed by daily traces that the high similarity of user mobility occurs between each day of a week. An interesting phenomenon is that *Friday* has a lower similarity than other days, since people always have fun on Friday night. The results verify that users in MSNets always keep relatively stable steady-state probability distributions over geocommunities, with occasional short-term fluctuations.

5 DESIGNING ALGORITHMS FOR THE SUPERUSER ROUTE

In this section, we investigate how the superuser controls its trajectory to quickly meet mobile users in the network. Since we have the knowledge of the geocentrality $C_i(t_i)$ for each geocommunity i in the network, our next step is to choose waiting times $t_i (\geq 0)$ at each geocommunity i , and order these geocommunities together to form a tour.

5.1 Static Route Algorithm

The total duration T of the superuser route has two components: 1) waiting time T_w : The sum of waiting times at the chosen geocommunities; 2) traveling time T_t : the total time that the superuser spends traveling between geocommunities. The total route time $T = T_w + T_t$. Assuming that only T_w contributes to the superuser's data dissemination, the superuser route design algorithm can be broken into two subproblems: finding a good set of geocommunities and their corresponding waiting times, and ordering these geocommunities together to form a tour.

5.1.1 Time-Sensitive Superuser

min-T-SRA. In this algorithm, the superuser aims to minimize the total duration of the route T under the constraint of a certain dissemination ratio. We can look at the two steps, "choosing appropriate geocommunities" and "constructing a path through them," independently. Since we have the knowledge of the geocentrality function of geocommunity $i (1 \leq i \leq J)$, our next step is to choose waiting times t_i corresponding to each geocommunity i , so that the total dissemination ratio for the superuser approaches p . Obviously, the geocommunities with $t_i \neq 0$ are selected as the stopping sites of the superuser. Clearly,

we want to minimize the total waiting time. The corresponding optimization problem is as follows:

$$\begin{aligned} \min \quad & \sum_{i=1}^J t_i \\ \text{s.t.} \quad & \sum_{i=1}^J C_i(t_i) \geq p \\ & t_i \geq 0, 1 \leq i \leq J. \end{aligned} \quad (3)$$

From (2), $C_i(t_i)$ is the sum of the logarithmic functions. Equation (3) then becomes a convex optimization problem, which can be solved by the methods presented in [18].

Once we have determined the geocommunities, we order them so as to minimize the length of the route. This amounts to the Traveling Salesman Problem (TSP), whose exact solution is NP-hard. TSP solvers like Concorde [19] can solve the problem accurately, for a few hundred points, within minutes. If the number of points is large, then we can choose any of the available approximation algorithms [20] that exist for TSP.

5.1.2 Dissemination-Ratio-Sensitive Superuser

max-p-SRA. In this case, the specific objective is to design a superuser route to maximize the dissemination ratio p before a certain deadline. Since geocentrality characterizes the *utility* of a geocommunity to the superuser's data dissemination, the optimization problem is accordingly given as follows:

$$\begin{aligned} \max p = \quad & \sum_{i=1}^J C_i(t_i) \\ \text{s.t.} \quad & \sum_{i=1}^J t_i + T_i(t_0, t_1, \dots, t_J) \leq T \\ & t_i \geq 0, 1 \leq i \leq J. \end{aligned} \quad (4)$$

If the constraint of (4) contains integer variance, the solution is NP-hard. However, this problem can be simplified to a knapsack problem. An intuitive approach [21] would be to consider the *geocentrality to traveling time ratio* e_i of each geocommunity, which is called the efficiency of a community with $e_i = \frac{C'_i(t)|_{t=0}}{t_{cur,i}}$, where $C'_i(0)$ indicates the gradient of $C_i(t_i)$ at $t_i = 0$, and $t_{cur,i}$ represents the traveling time from the current community to Community i . Next, one would select the communities with highest efficiency into the knapsack. The superuser can wait at each chosen community until the increment of geocentrality descends below a certain threshold δ . Obviously, selected communities generate the highest centrality while consuming the shortest total traveling time.

Note that a user can belong to several communities, which introduces a potential *overlap* among CUS of geocommunities in the network. However, the superuser route is comprised of some *ordered* geocommunities, i.e., in the form of geocommunity scheduling. It is possible that the superuser has already delivered the data to a certain user at one community. Then, the contribution from that user to the other associated communities should be excluded, because the superuser does not need to disseminate data to the same users more than once. That is why, we define the above algorithms for the superuser

route design as *Static Route Algorithm*, and further propose a *Greedy Adaptive Route Algorithm*. This introduces the scheme of updating geocentrality for communities at each step, in terms of all the noncontacted users in the network.

5.2 Greedy Adaptive Route Algorithm

In this algorithm, we also choose geocentrality as the community's utility, but instead, it computes geocentrality of *noncontacted users* for each community repeatedly.

Throughout the rest of this section, we use the following notation. Given a collection of geocommunities $\mathbb{S} = \{1, 2, \dots, J\}$ over a domain of users $\mathbb{M} = \{1, 2, \dots, N\}$. Let \mathbb{G} be a collection of *contacted* users (i.e., the users who have already received the data from the superuser). Assume $C_i(t_i)$ as the geocentrality function of geocommunity i during waiting time t_i , we further propose $\bar{C}_i(t_i)$ to denote such centrality of *noncontacted* users covered by geocommunity i (i.e., facing users not covered by set \mathbb{G}).

Algorithm 1 shows the details of GARA for *time-sensitive superuser* (i.e., *min-T-GARA*), where T represents the time constraint for the superuser route, and the subscript *cur* indicates the current community where the superuser stays. t_{soj} is the waiting time at the current community, and $t_{cur,j}$ indicates the traveling time from the current community to Community j , which is a constant and is known by the superuser as described before. $\bar{C}'_i(0)$ stands for the gradient of $\bar{C}_i(t_i)$ at $t_i = 0$. Note that *min-T-GARA* can be easily changed to *max-p-GARA* by modifying *Step 4* to the constraint of dissemination ratio.

Algorithm 1. Greedy Adaptive Route Algorithm for Time-sensitive Superuser (*min-T-GARA*)

- 1: $\mathbb{G} \leftarrow \emptyset; \mathbb{W} \leftarrow \mathbb{S}; T$
- 2: Compute $C'_i(0)$ for every $i \in \mathbb{S}$
- 3: Stop at the geocommunity with maximal $C'_i(0)$
- 4: **for** ($t = 1; t < T; t++$) **do**
- 5: **if** $User_k \in Community_{cur}$ **then**
- 6: $\mathbb{G} \leftarrow \mathbb{G} \cup User_k$
- 7: **end if**
- 8: $\mathbf{a}[i] = \bar{C}'_i(0), 1 \leq i \leq J, i \neq cur$
- 9: $temp = \mathbf{a}[j] = \max \mathbf{a}$
- 10: **if** $(\bar{C}'_{cur}(t_{soj}) \leq temp) \wedge (\bar{C}'_j(T - t - t_{cur,j}) \geq (\bar{C}'_{cur}(T) - \bar{C}'_{cur}(t_{soj})))$ **then**
- 11: Move to Community j
- 12: **else**
- 13: Stay at the current community
- 14: **end if**
- 15: **end for**

We elaborately illustrate *Step 8–11* in Algorithm 1. Intuitively, *GARA* aims to maximize the sum of centrality within the total duration of superuser route. Obviously, the superuser will choose the geocommunity with maximal $C'_i(t)|_{t=0}$ as the first stop. What matters is *if* and *when* the superuser should move to other geocommunities. Without loss of generality, we consider the condition of two geocommunities in the network. As shown in Fig. 8, suppose there are two geocommunities with $C'_1(t)|_{t=0} > C'_2(t)|_{t=0}$, and the traveling time $t_{1,2}$ between two geocommunities is a constant, given a superuser speed.

Assumptions on *when* and *if* the superuser should move to the other geocommunities are

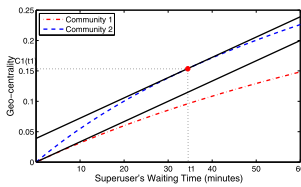


Fig. 8. Sketch illustrating the optimal transition time instant t_1 .

1. $C1$: The waiting time t_1 for the superuser staying at Geocommunity 1 before leaving for Geocommunity 2 is

$$C_1'(t)|_{t=t_1} = C_2'(t)|_{t=0}. \quad (5)$$

2. $C2$:

$$C_2(t_2) \geq C_1(T) - C_1(t_1). \quad (6)$$

Theorem 1. Suppose assumptions $C1$ - $C2$ hold, then the total centrality will achieve maximum within time constraint T .

$C2$ is obvious, since if the travel cost of moving to the other geocommunity is less than the total utility gain, the superuser should move; otherwise, the superuser would be better to stay at the current geocommunity. However, we prove the optimal transition time instant ((5)) in the Appendix, which can be found on the Computer Society Digital Library at <http://doi.ieeecomputersociety.org/10.1109/TPDS.2012.171>.

Note that the prerequisite of Theorem 1 is that the two geocommunities have unchanged centrality functions, whereas $GARA$ faces the dynamic centrality of geocommunities. However, the algorithm can guarantee the maximal total utility for the whole system at the transition time instant (i.e., t_i).

In contrast to SRA , $GARA$ can overcome the overlap among geocommunities in the network by facing noncontacted users each step, but the tradeoff is introducing more computational overhead.

6 PERFORMANCE EVALUATION

6.1 Simulation Setup

Our evaluations are conducted with Matlab on a realistic data set, *Infocom '06*, with AP locations on the map. We extract the distance between any two APs from the map of conference site,⁴ and treat it as the moving distance of the superuser between the two corresponding geocommunities. We compare our schemes (SRA and $GARA$)⁵ with the following two Message-Ferry-based routing schemes [8], [12] for *time-sensitive* ($min-T$) and *dissemination-ratio-sensitive* ($max-p$) superusers, respectively

- **Message Ferry moves with Restricted Random Waypoint model ($MF-RRWP$)**. The ferry moves according to the random waypoint mobility model, with the restriction that the waypoints are only chosen from the center of each geocommunity. At

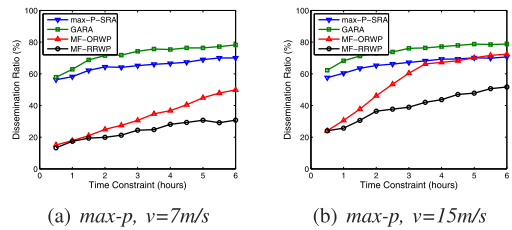


Fig. 9. Dissemination ratio under different superuser route schemes, and different superuser speeds.

each waypoint, the ferry pauses for exponentially distributed time with a mean of 15 minutes.

- **Message Ferry moves along ordered set of waypoints ($MF-ORWP$)**. The ferry orders the center of each geocommunity (waypoints) to form a shortest possible tour using the Concorde Traveling Salesman solver [19]. The ferry traverses the ordered set of waypoints repeatedly.

In our simulation, we focus on the following two metrics, which are key characteristics in data dissemination of MSNs:

- **Dissemination ratio**. The ratio of the number of delivered users to the total number of users in the network.
- **Average cost**. The traveling distance of the superuser. Note that although the superuser is not limited in power supply, we still aim to maximize the energy efficiency.

6.2 Performance Comparison

In general, $min-T-SRA$ and $max-p-SRA$ both select geocentrality as the utility of geocommunity; then they have almost the same performance improvement trend compared with other existing schemes in the respective application scenarios. Since *time-sensitive* case has been partially studied in [18], we pay particular attention to the analysis of *dissemination-ratio-sensitive* (i.e., $max-p$) case in this section.

Intuitively, a reader might think that the $MF-ORWP$ scheme would perform well, since it covers the entire region. However, there are good reasons for its poor performance. For time constraint $T = 2$ h and the superuser speed 15 m/s, the length of the route (tour) for $MF-ORWP$ is 150 km, compared with 40 km, which we observed for $GARA$ in the same setting. Fig. 10 shows the high superuser overhead of $MF-ORWP$, almost equal to $MF-RRWP$, and keeps two times SRA and three times $GARA$. With the same speed, the longer route length means that the superuser takes a longer time on the journey, and a significant fraction of this time is spent traveling in the parts of the region that have zero or negligible probability of user presence. Note that even though the superuser covers the entire region in $MF-ORWP$, it does not cover the entire region at once, especially when the region area is very large; therefore, the users and superuser can keep missing each other. We can also observe from Figs. 9a and 9b that the dissemination ratio of $MF-ORWP$ rises remarkably as the superuser speed increases from 7 to 15 m/s.

For $MF-RRWP$, the superuser may choose random geocommunities having no mobile user; thus, time spent traveling to, and staying at such communities, is completely futile. Therefore, $MF-RRWP$ performs almost the worst in terms of dissemination ratio and average cost.

4. The specific simulation parameters refer to [18].

5. We set 3 minutes as time unit in SRA and $GARA$.

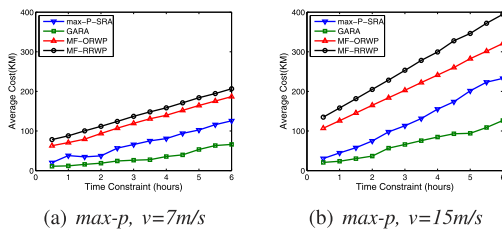


Fig. 10. Average cost under different superuser route schemes, and different superuser speeds.

Overall, *SRA* and *GARA* both perform significantly better than the other two MF-based schemes, with higher dissemination ratios, lower transmission delays, and less superuser overhead. The main reasons are that we balance the traveling time and waiting time. In addition, we invest waiting time at geocommunities that are more advantageous in terms of increasing the contact probability with the mobile users. It can be observed from Figs. 9a and 9b that *max-p-SRA* performs almost as well as *GARA* when the time constraint T is small (30 min in the experiment setting). This is because the superuser in *max-p-SRA* does not have enough time to visit a certain user more than once in such a short time frame. As T gets larger, *GARA* takes advantage of updating geocentrality metrics in terms of noncontacted users.

We also study the fairness of a user-superuser meeting among users in the Appendix, available in the online supplemental material.

7 RELATED WORK

In the context of intermittently connected networks, e.g., DTNs, Pocket Switched Networks (PSNs), and Opportunistic Networks, a number of routing schemes have been proposed for data forwarding [22], [5] and content dissemination [23], [24]. These routing schemes exploit the fact that end-to-end paths do exist over time in intermittently connected networks, which depend on a store-carry-and-forward pattern. Most recent work focuses on proposing routing schemes to achieve comparable performance as Epidemic routing, but with a lower cost measured by the number of relays needed for forwarding. Spray and wait [25] and its extended scheme Spray and focus [26] both select a fixed number of data relays, while some other schemes make relay selection decisions based on the nodes' data forwarding metrics. In [22], a relay forwards data to another node, whose forwarding metric is higher than itself. Delegation forwarding [27] is a single-copy forwarding scheme, which reduces the cost, by only forwarding data to the node with the highest metric. However, all of these schemes use the intrinsic mobility of the nodes in the network.

Another set of work considers the possibility of controlled mobility for network routing. They have proposed communication models where special mobile nodes (Message Ferry [8], [9], and Data MULEs [10], etc.) facilitate the network connectivity. However, these models always assume the special nodes move with fixed routes. SCFR [11] studies a multiple-ferry scenario, and the ferry trajectory is adaptive to the actual traffic and location of destinations. Moreover, multiple relays are allowed in SCFR, but with control. However, only ferries are mobile and all other nodes are static. Tariq et al. [12] aim toward designing a customized ferry route without disturbing

nodes' movements in mobile DTNs. However, they laid many constraints on node mobility, and did not consider the social nature of the network. On the contrary, our data dissemination scheme exploits the social characteristics of mobile networks without any online collaboration between the superuser and regular users in the network. Though we focus on a different application (data broadcasting from the superuser to regular users), our superuser also can extend to work as a "data carrier" between regular users. As such, it strengthens the research of both mobility-assisted routing schemes, and even the foundations in the area of intermittently connected networks.

8 CONCLUSIONS

In this paper, we have studied one-hop data broadcasting from a single superuser to other users in MSNets. The main idea behinds this is exploring both geographic and social properties of users mobility to facilitate data dissemination on purpose. We explore the geographic and social regularities of users mobility from both theoretical and experimental perspectives. Based on the exploited characterization, we introduce a semi-Markov process for modeling users mobility. The proposed superuser route comprises several geocommunities and the according waiting times, which are both calculated carefully based on the semi-Markov model. Extensive trace-driven simulation results show that our data broadcast schemes perform significantly better than other existing schemes. We also discuss multiple superuser scheduling and incentive schemes in selfish MSNets in the Appendix, available in the online supplemental material.

ACKNOWLEDGMENTS

A preliminary version of this paper was presented at IEEE MASS 2010, CA, 8-11 November 2010, and in conference proceedings [18]. This research was supported in part by NSFC 61222305, NSFZJ 1100324, SRFDP 20120101110139, and 863 2011AA040101-1. Jiming Chen is the corresponding author.

REFERENCES

- [1] S. Ioannidis, A. Chaintreau, and L. Massoulié, "Optimal and Scalable Distribution of Content Updates over a Mobile Social Network," *Proc. IEEE INFOCOM*, 2009.
- [2] B. Han, P. Hui, V.S.A. Kumar, M.V. Marathe, and G. Pei, "Cellular Traffic Offloading through Opportunistic Communications: A Case Study," *Proc. ACM Workshop Challenged Networks (CHANTS)*, 2010.
- [3] J. Wu and F. Dai, "Efficient Broadcasting with Guaranteed Coverage in Mobile Ad Hoc Networks," *IEEE Trans. Mobile Computing*, vol. 4, no. 3, pp. 1-12, May/June 2005.
- [4] S. Yang and J. Wu, "Efficient Broadcasting Using Network Coding and Directional Antennas in MANETs," *IEEE Trans. Parallel and Distributed Systems*, vol. 21, no. 2, pp. 148-161, Feb. 2010.
- [5] P. Hui, J. Crowcroft, and E. Yoneki, "Bubble Rap: Social-Based Forwarding in Delay Tolerant Networks," *Proc. ACM MobiHoc*, 2008.
- [6] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in Delay Tolerant Networks: A Social Network Perspective," *Proc. ACM MobiHoc*, 2009.
- [7] F.J. Ros, P.M. Ruiz, and I. Stojmenovic, "Acknowledgment-Based Broadcast Protocol for Reliable and Efficient Data Dissemination in Vehicular Ad-Hoc Networks," *IEEE Trans. Mobile Computing*, vol. 11, no. 1, pp. 33-46, Jan. 2012.

- [8] W. Zhao, M. Ammar, and E. Zegura, "A Message Ferrying Approach for Data Delivery in Sparse Mobile Ad Hoc Networks," *Proc. ACM MobiHoc*, 2004.
- [9] W. Zhao, M. Ammar, and E. Zegura, "Controlling the Mobility of Multiple Data Transport Ferries in a Delay-Tolerant Network," *Proc. IEEE INFOCOM*, 2005.
- [10] R. Shah, S. Roy, S. Jain, and W. Brunette, "Data MULEs: Modeling a Three-Tier Architecture for Sparse Sensor Networks," *Elsevier Ad Hoc Networks J.*, vol. 1, pp. 215-233, Sept. 2003.
- [11] J. Wu, S. Yang, and F. Dai, "Logarithmic Store-Carry-Forward Routing in Mobile Ad Hoc Networks," *IEEE Trans. Parallel and Distributed Systems*, vol. 18, no. 6, pp. 735-748, June 2007.
- [12] M. Tariq, M. Ammar, and E. Zegura, "Message Ferry Route Design for Sparse Ad Hoc Networks with Mobile Nodes," *Proc. ACM MobiHoc*, 2006.
- [13] Q. Yuan, I. Cardei, and J. Wu, "Predict and Relay: An Efficient Routing in Disruption-Tolerant Networks," *Proc. ACM MobiHoc*, 2009.
- [14] J. Scott, R. Gass, J. Crowcroft, P. Hui, C. Diot, and A. Chaintreau, "CRAWDAD Data Set Cambridge/Haggle (v. 2009-05-29)," <http://crawdad.cs.dartmouth.edu/cambridge/haggle>, May 2009.
- [15] N. Eagle, A. Pentland, and D. Lazer, "Inferring Social Network Structure Using Mobile Phone Data," *Proc. Nat'l Academy of Sciences*, vol. 106, no. 36, pp. 15274-15278, 2009.
- [16] A. Peddemors, H. Eertink, and I. Niemegeers, "CoSphere Data Set," <http://crawdad.cs.dartmouth.edu/novay/cosphere>, May 2008.
- [17] P.N. Tan et al., *Introduction to Data Mining*. Pearson, 2006.
- [18] J. Fan, Y. Du, W. Gao, J. Chen, and Y. Sun, "Geography-Aware Active Data Dissemination in Mobile Social Networks," *Proc. IEEE Int'l Conf. Mobile Adhoc and Sensor System (MASS)*, 2010.
- [19] W. Cook, D. Applegate, R. Bixby, and V. Chvatal, "Concorde: A Code for Solution of Travelling Salesman Problem," <http://www.tsp.gatech.edu/>, 2012.
- [20] V. Vazirani, *Approximation Algorithms*. Springer, Aug. 2001.
- [21] H. Kellerer, U. Pferschy, and D. Pisinger, *Knapsack Problems*. Springer, 2004.
- [22] E. Daly and M. Haahr, "Social Network Analysis for Routing in Disconnected Delay-Tolerant MANETs," *Proc. ACM MobiHoc*, 2007.
- [23] E. Yoneki, P. Hui, S. Chan, and J. Crowcroft, "A Socio-Aware Overlay for Publish/Subscribe Communication in Delay Tolerant Networks," *Proc. ACM Int'l Symp. Modeling, Analysis and Simulation of Wireless and Mobile Systems (MSWiM)*, 2007.
- [24] P. Costa, C. Mascolo, M. Musolesi, and G.P. Picco, "Socially-Aware Routing for Publish-Subscribe in Delay-Tolerant Mobile Ad Hoc Networks," *IEEE J. Selected Areas in Comm.*, vol. 26, no. 5, pp. 748-760, May 2008.
- [25] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and Wait: An Efficient Routing Scheme for Intermittently Connected Mobile Networks," *Proc. ACM SIGCOMM*, 2005.
- [26] T. Spyropoulos, K. Psounis, and C. Raghavendra, "Spray and Focus: Efficient Mobility-Assisted Routing for Heterogeneous and Correlated Mobility," *Proc. IEEE Int'l Conf. Pervasive Computing and Comm. Workshops*, 2007.
- [27] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, "Delegation Forwarding," *Proc. ACM MobiHoc*, 2008.



Jialu Fan received the BE degree in automation from Northeastern University in 2006, and the PhD degree in control science and engineering from Zhejiang University in 2011. She was a visiting scholar at Pennsylvania State University during 2009-2010. Currently, she is an assistant professor in the State Key Laboratory of Synthetical Automation for Process Industries at Northeastern University. Her research interests include mobile social networks, delay-

tolerant networks and networked control systems. She is a member of the IEEE.



Jiming Chen (M'08-SM'11) is a full professor in Department of control science and engineering, and the coordinator of group of Networked Sensing and Control in the State Key laboratory of Industrial Control Technology at Zhejiang University, China. His research interests are estimation and control over sensor network, sensor and actuator network, target tracking in sensor networks, optimization in mobile sensor network. He is a senior member of the IEEE.



Yuan Du received the BSc degree in control science and engineering from Zhejiang University in 2009. He is currently working toward the master's degree in the Department of Control Science and Engineering, and a member of the Group of Networked Sensing and Control in the State Key Laboratory of Industrial Control Technology at Zhejiang University. His research interests include the area of wireless sensor networks and mobile social networks.



Wei Gao received the BE degree in electrical engineering from the University of Science and Technology of China in 2005, and the PhD degree in computer science from the Pennsylvania State University in 2012. Currently, he is an assistant professor in the Department of Electrical Engineering and Computer Science at the University of Tennessee, Knoxville. He is a student member of the IEEE.



Jie Wu is the chair and a Laura H. Carnell Professor in the Department of Computer and Information Sciences at Temple University. Prior to joining Temple University, he was a program director at the NSF and Distinguished Professor at Florida Atlantic University. His research interests include wireless networks, mobile computing, and routing protocols. His publications include more than 600 papers in journals, conference proceedings, and books. He has served on several editorial boards, including TOC and JPDC. He was general co-chair for MASS 2006, IPDPS 2008, and DCOSS 2009 and was the program co-chair for INFOCOM 2011. He will be a general chair for ICDCS 2013. He was an IEEE Computer Society Distinguished Visitor and the chair for the IEEE Technical Committee on Distributed Processing. Currently, he is an ACM Distinguished Speaker. He received the 2011 China Computer Federation (CCF) Overseas Outstanding Achievement Award. He is a fellow of the IEEE.



Youxian Sun has been a full professor in Zhejiang University since 1988. He was an Alexander Von Humboldt Research Fellow at University of Stuttgart, Germany. In 1995, he was promoted to an Academician of the Chinese Academy of Engineering. His current research interests include modeling, control, optimization of complex systems, and robust control and its application. He is the president of the Chinese Association of Automation.

► For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.