# Geo-Social: Routing with Location and Social Metrics in Mobile Opportunistic Networks

Zhu Ying* Chao Zhang* Fan Li† Yu Wang*‡
* Department of Computer Science, University of North Carolina at Charlotte, Charlotte, NC 28223, USA
† School of Computer Science, Beijing Institute of Technology, Beijing, 100081, China.
‡ College of Information Engineering, Taiyuan University of Technology, Taiyuan, Shanxi, 030024, China.

***Abstract*—Mobile opportunistic networks (MONs) are intermittently connected networks, in which a multitude of mobile devices are carried by people and packets are delivered among devices via opportunistic communications. Routing in MONs is very challenging as it must handle network partitioning, long delays, and dynamic topology. Recently, new possibilities of social-based approaches which use social characteristics of mobile nodes to make forwarding decisions become a new trend in MONs. In this paper, we consider the location history with access patterns of a mobile user as its social features as well and propose several new *geo-social* metrics which reflect the location and social relationships among users. Several new routing algorithms are designed based on these new geo-social metrics to achieve efficient and stable routing in MONs. We evaluate them with a large-scale real-life mobile tracing dateset. Simulation results confirm the effectiveness of proposed geo-social methods.**

## I. Introduction

Mobile opportunistic network (MON) has recently drawn much attention from networking researchers due to its wide applications, such as mobile social networks, mobile data sharing, and participatory sensing. MON is a type of delay tolerant network (DTN). In MONs, a multitude of mobile devices are used and carried by people. The opportunistic communications allows delivering packets during intermittent contacts among mobile devices. However, the lack of instantaneous end-to-end paths, large transmission delay and unstable network topology make routing tasks in MONs very challenging.

Recently, the consideration of social characteristics of mobile users provides a new trend in the design of routing protocols for MONs: social-based approaches [1]. The mobility of mobile users heavily depends on their social characteristics such as carriers' social relations and behaviors. The social characteristics usually last for long term, and are more reliable in MONs environment, thus using them to make smarter forwarding decisions may significantly reduce the control overhead and improve routing performance. For example, *SimBet* [2] prefers a relay node with high social centrality and more common neighbors with the destination; *Label* [3] and *Group* [4] try to forward packets to a node within the same social group of the destination; while *Bubble Rap* [5] forwards data via a hierarchical community structure and chooses the node with higher centrality in the community.

Meanwhile, with the availability of enriched location information, location-based approaches have been used for designing communication protocols in different mobile networks. The most notable result is position based routing [6] in mobile ad hoc networks or sensor networks, where routing decisions are purely made based on the position information. There are mainly two ways to use location information in MON or DTN routing. Some DTN routing algorithms use the current location information to pick the next-hop relay node as the traditional position-based routing does. For example, GeoDTN+Nav method [7] combines the GPSR method [8] with DTN routing for a vehicular DTN. Packets are routed to the neighbor who has the smallest distance to the destination and if the current node does not have any neighbor at this moment, it switches to DTN mode by carrying the packet. The other set of location-based approaches [9]–[12] perform the mobility prediction via mobility pattern modeling over historical location information of nodes. For example, Leguay *et al.* [9] built a high-dimensional Euclidean space based on node location patterns. For each node, its coordinates are corresponding to its probability of being found in each location. By defining different distances between two nodes to represent their location similarity, several location-based DTN routings are proposed.

Although both social characteristics and location patterns of mobile users have already been proved effective in the design of MON routing, all existing approaches consider them isolated from each other. However, we believe that these two types of information could strongly related to each other in many scenarios. For example, the location history of an individual could imply, to some extent, his/her social interests and behaviors. People who share similar location histories are likely to have common interests, behaviors and some kind of relations. Family members are living at the same place (home) and colleagues stay in the same office during the day; classmates who take same classes may have the similar time and spacial schedule; and members of the table tennis club always show up at the same time in the gym. Therefore, it is possible to analyze the enriched location information and extract location/social characteristics among users. The new social characteristics, which are extracted from location

The work of Z. Ying, C. Zhang and Y. Wang is supported in part by the US National Science Foundation under Grant No. CNS-1319915 and CNS-1343355, and the National Natural Science Foundation of China under Grant No. 61428203. The work of F. Li is partially supported by the National Natural Science Foundation of China under Grant No. 61370192 and 61432015. The authors would like to thank Orange and the D4D challenge organizers for providing them the D4D datasets and allowing them to continue working on the datasets after the D4D challenge.

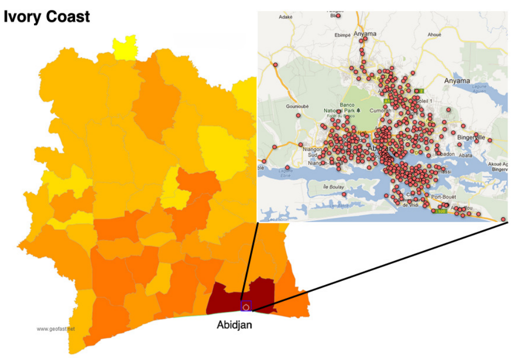



Fig. 1. **Visualization of D4D dataset**: the traffic distribution among all regions are shown in color where darker color indicates heavier traffic loads; the small blue rectangle around Abidjan shows the limited region with high traffic that we select for our simulations; and a zoomed view of the detailed tower distribution within the selected region is also provided.

information, can more accurately reflect the physic contact opportunities among users. We call them *geo-social* metrics. By exploring such kind of social and spatial characteristics, we propose several new geo-social based routing algorithms for large-scale MONs. Simulation results based on a real-life tracing dataset of large-scale mobile users demonstrate the efficiency of the proposed methods. compared with existing DTN routing methods.

The rest of this paper is organized as follows. We first briefly introduce the real-life mobile tracing dataset we used in Section II. Then we introduce two simple location-based methods in Section III. A set of geo-social characteristics and relative routing algorithms are designed and presented in Section IV. Section V presents our simulation results. Finally, Section VI concludes this paper.

## II. D4D Datasets

The appearance of mobile devices equipped with sensors (especially GPS) and contact/event logs enable pervasive monitoring of mobile users/devices' behaviors and mobility. Recently, there are several cellular datasets collected via smartphone based testbeds, such as Nokia Data Collection Campaign [13], MIT Reality [14], and D4D Challenge [15]. These real-life tracing data provide abundant resources to study social, spatial, and temporal characteristics of mobile users in different environments. To simulate large scale MONs, we choose the largest dataset that we can obtain, the D4D dataset, which is a real life wireless tracing data from the cellular operator Orange for the *Data for Development (D4D) challenge* [16]. The released D4D datasets [15] are based on anonymized Call Detail Records (CDR) of phone calls and SMS exchanges between $50,000$ Orange mobile users in Ivory Coast between December 1, 2011 and April 28, 2012 (150 days). We use the dataset of individual trajectories with high spatial resolution (*SET2* in D4D datasets), which contains the access records of cellular tower of each mobile user over every two-week period during the whole duration of D4D data collection period. We use the sequences of visited cellular towers of all users to generate both contact encounters among mobile users and location/social profiles of each user.

Since D4D datasets do not have direct encounter information between phones via short range communications (such as Bluetooth or WiFi), to support opportunistic communications we assume that two phones can direct communicate with each other if they share the same cellular tower at a particular time. Though this assumption may not be true in reality, it gives us an approximated environment for opportunistic communications in such a large scale network. We use all access records for every two-week periods, but only perform the opportunistic communications within each two-week period. For one two-week period, there are already huge number of users and encounters. For example, in the first two-week period, there are $46,254$ active mobile users, $1,097$ cellular towers, and $6,787,594$ encounters between users in total. In a previous study [17], we have shown that the smaller size of user set and encounter database could accelerate the execution time of our simulations while conclusions from simulation results are still consistent with those using the whole user set. Therefore, in this paper, we choose a subset of users from the whole user set by requiring that the users must be from the first 15,000 users in our encounter database and the physical locations of encounters must be within a small region. Fig. 1 shows the number of calls during a period of two weeks (where darker color indicates heavier traffic loads). Clearly, the traffic load distribution within Ivory Coast is unbalanced. Therefore, we choose a small region (a blue rectangle region in Fig. 1, whose longitude and latitude range from $[-8.49, -2.69]$ and $[4.41, 10.47]$, respectively) with the heaviest traffic load. This region is around Abidjan, the economic and former official capital of Ivory Coast and the largest city in the nation. A zoomed view of this region in Fig. 1 shows that it holds a large number of cellular towers. For the first two-week period, this selected dataset has $6,318$ active mobile users, $496$ cellular towers, and $327,717$ encounters between users.

## III. Location Based Routing

In this section, we introduce two simple location-based routing methods, which mine the similarity between users based on their geographic location histories (i.e., cell tower ID scan records).

### A. Location Based Routing

A user's visit frequency and duration of a place in the past may imply the possibility of this user to visit this place in the future. Therefore, by recording the historical visiting frequency and duration of each location, we can build a *location profile* of each user, which reflects how likely this user will visit a particular place. In this paper, we use the location of each cellular tower as one location. However, our proposed method can work with other definition of locations. Assume that there are $n$ mobile users $(v_1, \cdots, v_n)$ and $m$ cellular towers $(t_1, \cdots, t_m)$. Then we can define the location profile of a user $v_i$ as follows:

*Definition 1 (Location Profile):* The location profile of user $v_i$ is defined as a $m$-dimensional vector,

$$L(v_i) = \{p(v_i, t_1), \cdots, p(v_i, t_m)\},$$

where $p(v_i, t_j) = \frac{d_{ij}}{\sum_{j=1}^{m} d_{ij}} \cdot \frac{f_{ij}}{\sum_{j=1}^{m} f_{ij}}$. Here $d_{ij}$ and $f_{ij}$ are the total visiting duration and frequency of user $v_i$ to tower $t_j$, repsectively. Thus, $p(v_i, t_j)$ basically shows the product of the portion of duration/frequency of user $v_i$ to tower $t_j$ compared with the total duration/frequency to all towers. Larger value of $p(v_i, t_j)$ generally indicates $v_i$ visiting $t_j$ more often and staying longer.

Based on location profiles, we can calculate the location similarity of two users as follows:

*Definition 2 (Location Similarity):* The Geo-Similarity of two users $v_i$ and $v_j$ is defined as the inner product of their location profiles, i.e,

$$Sim(v_i, v_j) = L(v_i) \cdot L(v_j) = \sum_{x=1}^{m} p(v_i, t_x) p(v_j, t_x).$$

Location similarity implies the similarity between location visiting patterns of these two users, which hopefully reflects the probability of their meeting at a cell tower in the future. With the definition of location similarity, the location based routing method is straightforward.

- **Location (Loc):** The message is only forwarded from $v_i$ to $v_j$ during an encounter between them if and only if $v_j$ has higher location similarity with the destination $v_d$ than $v_i$ does (i.e. $Sim(v_j, v_d) > Sim(v_i, v_d)$). Similar ideas have been used in [9].

### B. Top-Location Based Routing

Notice that each node needs to compute and track its own location profile while the source node needs to obtain the destination's location profile and put it into the message head. If the network spans a large number of places (i.e. the number of towers $m$ is large), the size of location profiles could be huge which lead to large number of overhead (due to both long messages and large storage requests). Therefore, we try to simplify the calculation of location profile and location similarity among nodes.

From the observation of *location profiles* of mobile users in D4D dataset, we find that the towers which users visit with a large frequency and long durations are mainly limited to a small number of towers. Table I gives an example of a ordered location profile for one user in D4D Dataset. Therefore, we limit the number of locations within the profie to those top-$k$ visited places instead of every towers to describe the user's location profile. In addition, from Table I we can find that the visiting duration of a tower is directly proportional to its visiting frequency in most of the cases. Plus most of the encounter durations are already long enough to perform a message exchange, so we simply use the most frequently visited $k$ locations of a user as the representation of the user's location characteristics. In this paper, we set the default value of $k$ as 10. Each node $v_i$ thus only needs to maintain a top-10 place list with highest visiting frequencies, denoted as $TP(v_i)$. Then the similarity between two users $v_i$ and $v_j$ can be estimated as the *number of common top places* (i.e. $ComTP(v_i, v_j) = |TP(v_i) \cap TP(v_j)|$). The *top-location based routing* works as follow:

TABLE I
SORTED LOCATION PROFILE OF A D4D USER

| Tower ID | 750 | 1129 | 953 | 898 | 303 | 163 | 1022 | $\cdots$ |
|---|---|---|---|---|---|---|---|---|
| Duration | 33,720 | 25,354 | 15,723 | 4,268 | 210 | 113 | 36 | $\cdots$ |
| Frequency | 216 | 187 | 135 | 52 | 8 | 5 | 2 | $\cdots$ |

- **Top Location (TopLoc):** a message is only forwarded from $v_i$ to $v_j$ if and only if $v_j$ has more common top locations with the destination $v_d$ than $v_i$ has (i.e. $ComTP(v_j, v_d) > ComTP(v_i, v_d)$).

## IV. GEO-SOCIAL BASED ROUTING

In this section, we propose some new geo-social routing methods which jointly consider the location information and social properties of mobile users. First, we introduce how to extract user's social properties from their location information. Then we show how to use these properties and/or other social/location characteristics as routing metrics.

### A. Geo-Social Metrics: Home and Work Places

The users' location information can reflect their social relationships. For example if two users' homes are very close, they are neighbors. If they live at the same place, they are probably family members. If the workplaces of two users are at the same place, they might be colleagues. Most people spend a huge percentage of their time at their home and work places, (for example, a normal person who sleeps for eight hours a day and works for eight hours during weekdays, thus he/she will stay at home or workplace for up to $57\%$ of time in his/her life). Therefore, we believe that where a people live and work represents the people's most important social features and detecting such locations can help us to understand regular communication opportunities.

Based on our recent study [18], we can use the following simple methods to detect a user's home and workplace. We let the top $k$ towers (in term of total frequency) that a user visits during "sleeping time" (11pm$\sim$7am) as the home places of this user (denoted by $H(v_i)$). Similarly, the top $k$ towers that a user visits during daytime (8am$\sim$7pm) of weekdays as the workplaces of this user (denoted by $W(v_i)$). In [18], we have shown that such simple rule based method can achieve over $70\%$ accuracy over Nokia Data Collection Campaign dataset [13]. In this paper, we set $5$ as the default value of $k$ so that in total we maintain $10$ home/work places per node.

With home/work places to describe an individual's living and working features, we can also define two new metrics to approximatively measure how likely two nodes will be colleagues or neighbors: the *number of common home places* (i.e. $ComH(v_i, v_j) = |H(v_i) \cap H(v_j)|$) and the *number of common workplaces* (i.e. $ComW(v_i, v_j) = |W(v_i) \cap W(v_j)|$). Clearly, larger number of common home/work places indicates stronger social tie between two users. With this social information, we then can define a *user-home* graph $G_h$ where there is an edge between two users $v_i$ and $v_j$ if and only if they share at least $\alpha$ common home places, i.e., $ComH(v_i, v_j) \geq \alpha$. Similarly, a *user-work* graph $G_w$ can be defined. Both $G_h$ and $G_w$ are social graphs among all mobile users. In this paper, we let $\alpha = 2$. We can calculate any social properties (such as degree centrality or community structures) over these graphs. Hereafter, we use $c_h(v_i)$ and $c_w(v_i)$ denote the degree centralities of a user $v_i$ over $G_h$ and $G_w$, respectively.

### *B. Geo-Social Based Routing*

In this section, we propose five geo-social routing methods.

*1) Geo-Social Based Routing Based on Work/Home:* The first two methods use the numbers of common home and work places. Naturally, the more common home/work places between two users, the more likely they live/work in the same area. Thus the relay node is selected by comparing $ComW(v_j, v_d)$ and $ComH(v_j, v_d)$ of the neighbor $v_j$ with $ComW(v_i, v_d)$ and $ComH(v_i, v_d)$ of the current node $v_i$.

- **Work:** a message is only forwarded from $v_i$ to $v_j$ if and only if (iff) $v_j$ has more common work places with $v_d$ than $v_i$ has, i.e. $ComW(v_j, v_d) > ComW(v_i, v_d)$.
- **Work/Home:** a message is only forwarded from $v_i$ to $v_j$ iff $ComW(v_j, v_d) + ComH(v_j, v_d) > ComW(v_i, v_d) + ComH(v_i, v_d)$ or $ComW(v_j, v_d) > ComW(v_i, v_d)$.

Note that we can also define a similar method purely based on $ComH(v_i, v_d)$. However, people usually have more contact opportunity at work place than at home.

*2) Geo-Social Based Routing with Multiple Metrics:* Using multiply metrics as relay selection criterion may benefit by taking advantage of varies information resources. We now explore the combination of social- and location-based metrics in the following three geo-social based routing methods.

- **Location+Centrality (Loc+Centr):** a message is only forwarded from $v_i$ to $v_j$ if and only if $Sim(v_j, v_d) > Sim(v_i, v_d)$ or $c(v_j) > c(v_i)$. Here $c(v_i)$ is the degree centrality from the contact graph $G$ where two nodes share an edge if they have met before.
- **Work/Home+Centrality (W/H+Centr):** a message is only forwarded from $v_i$ to $v_j$ iff $ComW(v_j, v_d) > ComW(v_i, v_d)$ or $ComW(v_j, v_d) + ComH(v_j, v_d) > ComW(v_i, v_d) + ComH(v_i, v_d)$ or $c(v_j) > c(v_i)$.
- **Work/Home+Bubble (W/H+Bubb):** inspired by [5], we jointly consider both common home/work places with the destination ($ComH(v_i, v_d)$ and $ComW(v_i, v_d)$) and the home/work centrality ($c_h(v_i)$ and $c_w(v_i)$) in home/work social graphs. When the current node doesn't have any common *home/work places* with the destination, it chooses the relay node with higher *home/work centrality* or with common *home/work places* with $v_d$ since it is not close to either home or work places of the destination. Until the current node meets with a node, which has common *home/work places* with the destination, it then only relays the message to the node, with more *common home/work places* with the destination (same as **Work/Home**). Detail is given by Algorithm 1.

**Algorithm 1** Work/Home+Bubble Routing (W/H+Bubb)

When node $v_i$ with a message $M$ destined to $v_d$ encounters a node $v_j$.

```
if v_j = v_d then
    v_i forwards M to v_j
else if ComW(v_i, v_d) + ComH(v_i, v_d) = 0 then
    if c_h(v_j) > c_h(v_i) or c_w(v_j) > c_w(v_i) or
       ComW(v_j, v_d) > 1 or ComH(v_j, v_d) > 1 then
        v_i forwards M to v_j
    else
        v_i holds on M
else
    if ComW(v_j, v_d)+ComH(v_j, v_d) > ComW(v_i, v_d)+
       ComH(v_i, v_d) or ComW(v_j, v_d) > ComW(v_i, v_d)
    then
        v_i forwards M to v_j
    else
        v_i holds on M
```

## V. SIMULATIONS

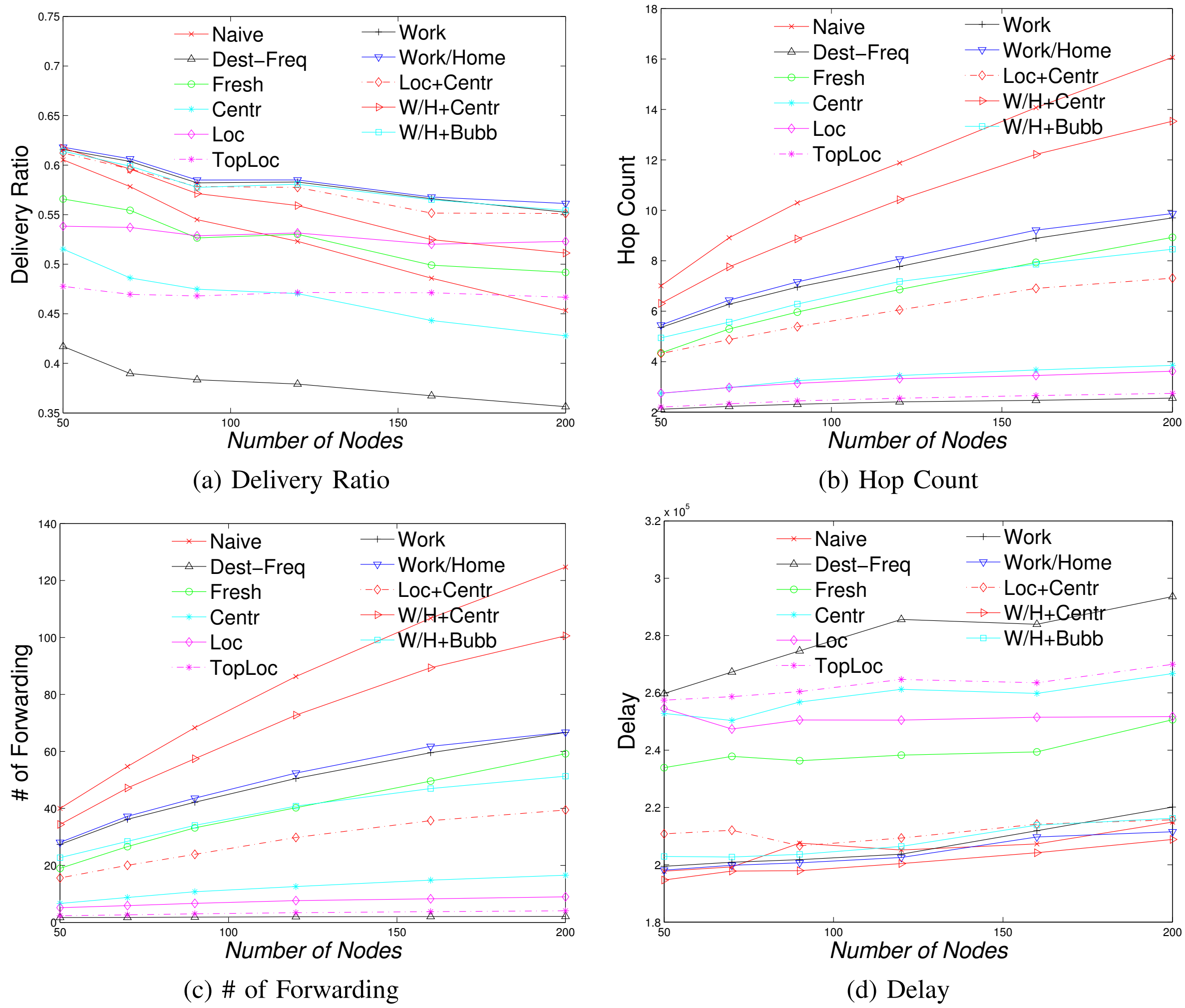
To test our proposed geo-social based routing methods, we also implement four representative MON routing algorithms.

- **Naive:** the message is always forwarded to the encountering node during any encounter. It can be treated as a single-copy version of Spray and Wait [19].
- **Fresh:** [20] the message is only forwarded from the current node $v_i$ to the encountered node $v_j$ if $v_j$ has met the destination more recently than $v_i$ does.
- **Destination Frequency (Dest-Freq):** [21] the message is only forwarded from $v_i$ to $v_j$ if $v_j$ has met the destination more often than $v_i$ does.
- **Centrality (Centr):** the message is only forwarded from $v_i$ to $v_j$ if $v_j$ has higher centrality than $v_i$ does. Here, we use the degree centrality as the centrality metric, similar to many existing social-based methods [2], [5].

Plus the seven routing methods introduced in this paper (Loc, TopLoc, Work, Work/Home, Loc+Centr, W/H+Centr, W/H+Bubb), we totally test eleven routing methods over the D4D data set described in Section II. We conduct extensive simulations by measuring the following four metrics:

- *Average delivery ratio*, the average percentage of successfully delivered messages from source to destination.
- *Average hop count:* the average number of hops during each successful delivery from source to destination.
- *Average number of forwarding:* the average number of messages forwarding in the network.
- *Average delay* (in secs), the average time duration of successfully delivered messages from source to destination.

For all experiments, we pick all possible source and destination pairs among $n$ active mobile users and perform $\frac{n(n+1)}{2}$ routing tasks among these selected participators in each two-week period. All results reported here are the average over these tasks and over different two-week periods. For all routing algorithms, we use their multiple copy versions where the




(a) Delivery Ratio

(b) Hop Count

(c) # of Forwarding

(d) Delay

Fig. 2. Performances of different geo-social routing algorithms with various numbers of nodes $n$.

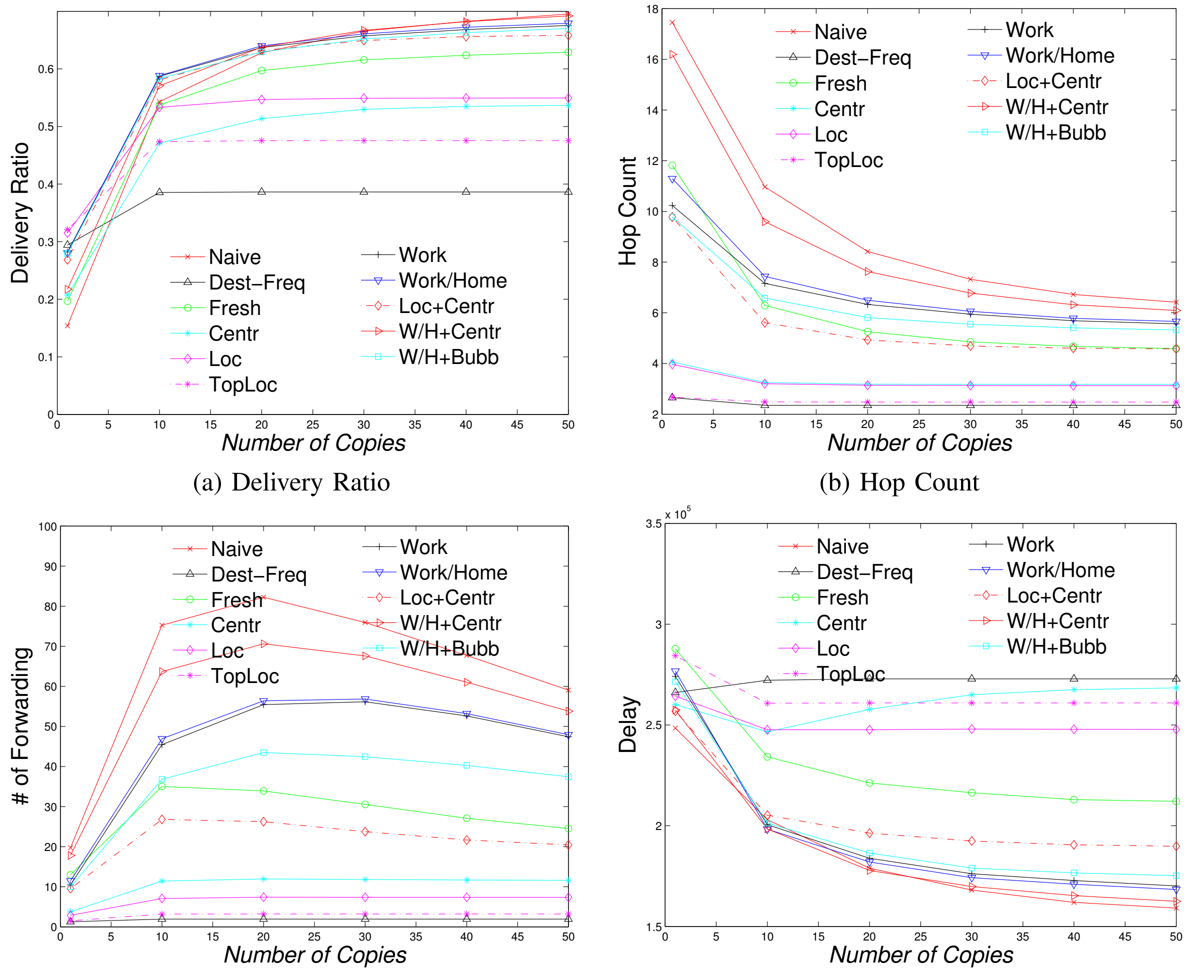
number of duplicates of a message on mobile users is limited by a constant $\beta$. Such maximum number of copies can be implemented using tokens as in [19].

In the first set of simulations, we fix the number of copies allowed $\beta$ at 10 and increase the number of participators $n$ from 50 to 200. Simulation results are reported in Fig. 2. With more participators, the performances are slightly reduced. This is reasonable, since it becomes more challenging to delivery packets among more mobile users, especially, when we always select $n$ most active users among all D4D users. From Fig. 2(a), we can observe the following facts: (1) Dest-Freq has the lowest delivery ratio, while several of our proposed geo-social methods (Work, Work/Home, Loc+Centr, W/H+Bubb) can achieve the highest delivery ratio. All of our geo-social methods can achieve better performances than methods purely using social or location information (such as Loc or Centr). (2) Compared with Loc, TopLoc has much worse performance. This is reasonable since less information is used. However, Work and Work/Home also maintain small location information but perform very well. This demonstrates the power of geo-social metrics which combines both location and social characteristics of mobile users. Finally, from the remaining figures in Fig. 2, geo-social based methods also have relatively smaller number of hops and shorter delay.

We also study the effect of the number of allowed replicas $\beta$. In this set of simulations, we fix the number of nodes at 100 and increase $\beta$ from 1 to 50. From Fig. 3, we can observe that larger number of replicas results in higher delivery ratio, smaller average hops and shorter delay in general. It is obviously that the number of forwarding gets larger with addition number of replicas. Note that after $\beta$ reaches around 20, some of the methods have smaller number of forwardings. This may due to the higher delivery ratios. Clearly, there is a tradeoff between routing performance and communication overheads. Among all methods, geo-social based methods can usually achieve better performances. One interesting discovery is that naive method can achieve better performance than some of the proposed methods when $\beta$ is large. This is because naive method becomes epidemic routing when $\beta$ is large, which shows the upper bound of delivery ratio of any routing method.

## VI. Conclusion

In this paper, we investigate how to design new geo-social based MON routing methods which consider both social and




(a) Delivery Ratio (b) Hop Count

(c) # of Forwarding (d) Delay

Fig. 3. Performances of different geo-social routing algorithms with various numbers of copies allowed $\beta$.

location information of mobile users. By leveraging the semantic meaning of a location and mining the historical location information, we propose several new geo-social metrics and design several new routing methods which use these metrics to pick relay nodes. Via extensive simulations over a real-life large-scale wireless tracing data, we show the efficiency of geo-social based approaches in MON routing.

## REFERENCES

[1] Y. Zhu, B. Xu, X. Shi, and Y. Wang, "A survey of social-based routing in delay tolerant networks: Positive and negative social effects," *IEEE Communication Survey and Tutorials*, vol. 15, no. 1, pp. 387–401, 2013.

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

[3] P. Hui and J. Crowcroft, "How small labels create big improvements," in *Proc. of Int'l Workshop on Intermittently Connected MANETs*, 2007.

[4] F. Li, L. Zhao, C. Zhang, J. Gao, and Y. Wang, "Routing with multi-level cross-community social groups in mobile opportunistic networks," *Personal and Ubiquitous Computing*, 18(2):385–396, 2014.

[5] P. Hui, J. Crowcroft, and E. Yonek, "Bubble Rap: Social-based forwarding in delay tolerant networks," in *Proc. of ACM MobiHoc*, 2008.

[6] M. Mauve, et al., "A survey on position-based routing in mobile ad hoc networks," *IEEE Network*, vol. 15, no. 6, pp. 30–39, 2001.

[7] M. Gerla, P.-C. Cheng, et al., "GeoDTN+Nav: Geographic dtn routing with navigator prediction for urban vehicular environments," *Mobile Networks and Applications*, vol. 15, no. 1, pp. 61–82, 2010.

[8] B. Karp and H. T. Kung, "GPSR: Greedy perimeter stateless routing for wireless networks," in *Proc. of ACM MobiCom*, 2000.

[9] J. Leguay, T. Friedman, and V. Conan, "DTN routing in a mobility pattern space," in *Proc. of ACM WDTN*, 2005.

[10] W. Gao and G. Cao, "Fine-grained mobility characterization: Steady and transient state behaviors," in *Proc. of ACM MobiHoc*, 2010.

[11] J. Wu, et al., "Homing spread: Community home-based multi-copy routing in mobile social networks," in *Proc. of IEEE INFOCOM*, 2013.

[12] J. Link, D. Schmitz, and K. Wehrle, "GeoDTN: Geographic routing in disruption tolerant networks," in *Proc. of IEEE GLOBECOM*, 2011.

[13] Mobile Data Challenge, http://research.nokia.com/page/12000, 2012.

[14] Reality mining project, MIT Media Lab, http://reality.media.mit.edu/

[15] V. D. Blondel, M. Esch, C. Chan, F. Clerot, P. Deville, E. Huens, F. Morlot, Z. Smoreda, and C. Ziemlicki, "Data for Development: The d4d challenge on mobile phone data," in *arXiv.1210.0137v2*, 2013.

[16] Data for Development (D4D) Challenge, http://www.d4d.orange.com.

[17] Y. Zhu, C. Zhang, and Y. Wang, "Mobile data delivery through opportunistic communications among cellular users: A case study for the d4d challenge," in *Proc. of NetMob*, 2013.

[18] Y. Zhu, Y. Sun, and Y. Wang, "Nokia mobile data challenge: Predicting semantic place and next place via mobile data," in *Proc. of Nokia Mobile Data Challenge Workshop*, 2012.

[19] T. Spyropoulos, et al., "Spray and wait: an efficient routing scheme for intermittently connected mobile networks," in *ACM WDTN*, 2005.

[20] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, "Age matters: efficient route discovery in mobile ad hoc networks using encounter ages," in *Proc. of ACM MobiHoc*, 2003.

[21] V. Erranmilli, M. Crovella, A. Chaintreau, and C. Diot, "Delegation forwarding," in *Proc. of ACM MobiHoc*, 2008.