Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 7, July 2014, pg.912 – 920

RESEARCH ARTICLE

Community-Aware Opportunistic Routing in Mobile Social Networks

Mr. Shankarling Umadi
MTech Scholar
VTU PG RC
Gulbarga-India
Shankarumadi111@gmail.com

Ms. Ranjana B Nadagoudar
Assistant Professor
VTU PG RC
Gulbarga-India
ranjanapriya8@gmail.com

Ms. Amareshwari Patil
Associate Professor
PDAEC
Gulbarga-India
aspatilpdacse@gmail.com

Abstract— Mobile social networks (MSNs) are a kind of delay tolerant network that consists of lots of mobile nodes with social characteristics. Recently, many social-aware algorithms have been proposed to address routing problems in MSNs. However, these algorithms tend to forward messages to the nodes with locally optimal social characteristics, and thus cannot achieve the optimal performance. In this paper, so proposed a distributed optimal Community-Aware Opportunistic Routing (CAOR) algorithm. Our main contributions are that we propose a home-aware community model, whereby we turn an MSN into a network that only includes community homes. We prove that, in the network of community homes, we still can compute the minimum expected delivery delays of nodes through a reverse Dijkstra algorithm and achieve the optimal opportunistic routing performance. Since the number of communities is far less than the number of nodes in magnitude, the Computational cost and maintenance cost of contact information are greatly reduced.

1. Introduction

Mobile social networks (MSNs) are a special kind of delay tolerant network (DTN), in which mobile users move around and communicate with each other via their Carried short-distance wireless communication devices. Typical MSNs include pocket switch networks, mobile vehicular networks, mobile sensor networks, etc [1].Recently, some social-aware routing algorithms that are based on social network analysis have been pro- posed, such as Bubble Rap [2], SimBet [3], and algorithms in [4–7], etc. Two key concepts in social network analysis are: (i) *community*, which is a group of people with social relations; (ii) *centrality*, which indicates the social relations between a node and other nodes in a community. Based on the two concepts, these algorithms detect the communities and compute the centrality value for each node.

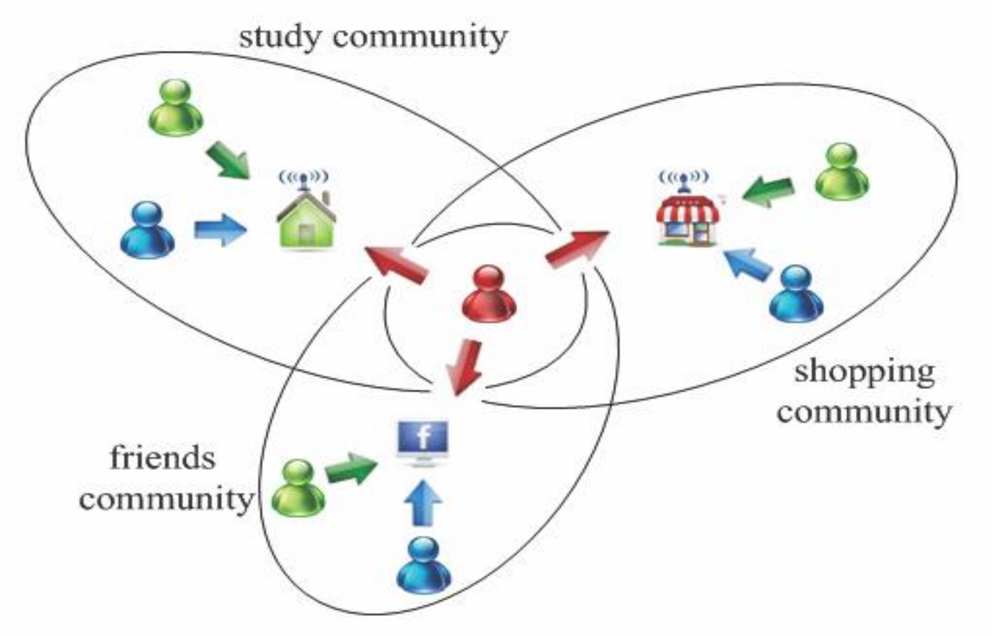


Fig.1.An example of mobile social network

In this paper, we focus on the single-copy routing problem in MSNs. In many real MSNs, mobile users that have a common interest generally will visit some (real or virtual) location that is related to this interest. For instance in Fig. 1, students with a common study interest will visit the same classrooms to take part in the same courses; customers with the same shopping interests often visit the same shops; friends generally share some resources through facebook, and so on. Based on this basic social characteristic, we propose a home-aware community model. Mobile users with the common interest autonomously form a community, in which the frequently visited location is their common "*home*". Moreover, like [1], we assume that each home supports a real or virtual throwbox [4], a local device that can temporarily store and transmit messages. Under the home-aware community model, we propose a distributed optimal Community-Aware Opportunistic Routing algorithm (CAOR). More specifically, our major contributions are summarized as follows:

1)We present a home-aware community model and extend the centrality concept from a single node to a group of nodes.

2) We present a rule of optimal opportunistic routing through a theoretical analysis. We design a reverse Dijkstra algorithm to determine the optimal relays and compute the minimum expected delivery delay.

3) We turn the routing in |V | mobile nodes into a routing in |L| (|L| «|V |) community homes by virtue of the home-aware community model.

2. NETWORK MODEL & ASSUMPTIONS

We consider an MSN composed of |V | nodes V ={v | v ∈ V } moving among |L| locations L={l | l ∈ L}(|L| « |V |). Each mobile node visits a few locations frequently, while visiting the others rarely. A typical MSN is the Wi-Fi campus network at Dartmouth College.

TABLE 1
Description of auxiliary variables

v, l, d, S	v is a node, l is a location or home, d is a destination, and S is a set of nodes or homes.
$\lambda_{v,l}, \lambda_{S,l}$	the exponential distribution parameter of that node v and (any node in) node set S visit home l.
$\tilde{R}_i$	the optimal relay set for the opportunistic routing from a message sender (i) to destination d.
$\tilde{S}$	the optimal betweenness set for the message delivery from home i to j. When the context is clear, the subscripts are removed.
$D_{i,d}(S)$	the minimum expected delivery delay from i to d via relay set S, where i may be a node, or a home. Specifically, $D_{i,d} = D_{i,d}(\tilde{R}_i)$ when $S = \tilde{R}_i$.
$B_{l,l'}(S)$	the betweenness of S, i.e., the expected delivery delay from home l to l' only via node set S.

Another MSN that follows this characteristic is the mobile vehicular network, in which lots of buses and taxies move among bus stations and taxi stops. In this paper, we call the frequently visited locations *homes*. Like in the previous work, we assume that the behavior of each mobile node visiting homes follows the Poisson process. In other words, the time interval that each node visits a home follows an exponential distribution (the unit of time is time slot).

We also assume that each home has a “throw- box” [4], which has the ability to store and transmit messages. In fact, many MSNs follow this assumption. For instance, the Road Side Units (RSUs) in mobile vehicular networks are a kind of real throwboxes. The APs in the Wi-Fi campus network can also be seen as a type of real throwboxes since each user can upload/download data from network storages via these APs.

3. SOCIAL NETWORK MODELLING

Before presenting our algorithm, we build the com- munity and define two social metrics.

3.1 Building Home-Aware Communities

In this paper, we propose a concept of home-aware community. A home-aware community is a community of nodes that frequently visit a given home

Definition 1: Home-Aware Community: a home-aware community C_l is a set of nodes that frequently visit home l (beyond a given threshold). That is:

$$C_l = \{v|\lambda_{v,l} \geq \varepsilon, v \in V\}. \quad (1)$$

Moreover, home l is equipped with a real or virtual throwbox, so that it can be used as a relay of message delivery

3.2 Centrality Metric

In an MSN, the centrality metric is generally used to measure the importance of nodes during message delivery. A node with a better centrality value means that it has a stronger capability of connecting with other nodes. Previous works mainly adopt three centrality measures: degree centrality, closeness centrality, and betweenness centrality [3]. Degree centrality is measured as the number of direct links between a given node and other nodes. Closeness centrality is a measure of how long it will take to deliver a message from a given node to other nodes.

A. Intra-Community Centrality

In intra-community routing, the most concern is measuring the capability of each community member to meet and deliver messages to other members. Since intra-community message deliveries happen only when nodes visit community homes, the smaller the expected delay to visit a community home, the higher capability to deliver messages a community member would have.

Definition 2: Intra-Community Centrality : $Il(v)$ is the reciprocal of the expected delay for node v visiting a community home l,
i.e., $Il(v)=1=Dv;l=\_v;l$.

B. Inter-Community Betweenness

In this paper, we adopt the opportunistic routing scheme, in which multiple nodes cooperatively deliver messages. It can be defined as follows.
Definition 3: Opportunistic Routing: each message sender (home or node) has a relay set (homes or nodes). Once a relay in the set meets the message sender, the sender will let this relay deliver messages. relay.

4. OVERVIEW OF CAOR

In this section, we introduce the methodology and basic idea of CAOR. Here, we assume that the source (and relays) knows which communities that the destination d belongs to. That is, the message consists of the source, the destination information, and the data to be delivered. For example, the destination can broadcast this information to all community homes, or a naming service is used to distribute this information, etc.

4.1 Methodology: Optimal Opportunistic Routing

The optimal opportunistic routing scheme means that each message sender delivers messages via its optimal relay set. The key problem is to determine whether a relay belongs to the optimal relay set for each message sender.
We consider an opportunistic routing from a message sender i to the destination d via some candidate relays $\{u|\_i;u>0\}$. Here, the message sender i might be a mobile node or a home. Each u is a one-hop relay of i, i.e.,$\_i;u > 0$, but it does not must be a one-hop relay of the destination. The optimal relay set, denoted by $\sim Ri$,is given by the following formula:

$$\sim Ri=argminS\subseteq\{u|\_i;u>0\}Di;d(S): \qquad (4)$$

In Eq.(4), $Di;d(S)$ is the expected delay for i delivering messages to d via the relay set S. Moreover, for simplicity, we let

$$Di;d=Di;d(\sim Ri): \qquad (5)$$

Theorem 2: Optimal Opportunistic Routing Rule:the message sender always delivers messages to the encountered relay that has a smaller minimum expected delay to the destination than itself. Concretely, a relay u belongs to the optimal relay set $\sim Ri$ for the delivery from i to d, if and only if,

$$Du;d<Di;d,\text{i.e.:}u\in\sim Ri\Longleftrightarrow Du;d<Di;d \qquad (6)$$

4.2 Optimality of CAOR

First, we show that CAOR can achieve the minimum expected delivery delay in the simplified network. As the description in Section 4.2, CAOR uses a reverse Dijkstra algorithm to calculate the minimum expected delivery delay from each home to the destination in the extended graph $G+$. The minimum expected delivery delays will be derived

in an ascending order. When we compute the (k+1)-th minimum expected delivery delay, the delays smaller than this one have been derived out.
According to the optimal opportunistic routing rule in Theorem 2, only these homes can be candidate relays of the home with the(k + 1)-th minimum expected delivery delay.

Second, we can get that the minimum expected delivery delays that are derived in the simplified network are equal to the corresponding values in the original network.

Theorem 3: Assume that community Cl has m overlapped communities $Cl1 ; \cdots ;Clm$.
Then, the optimal relay set $\sim Rl$ of home l, and the optimal betweenness sets $\sim Sl;li$ ($1\leq i\leq m$) satisfy:
1) if $v \notin \cup mi=1\sim Sl;li$, then $v \notin \sim Rl$;
2) $\sim Sl;li \subseteq \sim Rl$, otherwise $\sim Sl;li \cap \sim Rl=\emptyset$ for $\forall i\in[1;m]$.
Theorem 3 shows that either all nodes in the optimal betweenness set $\sim Sl;li$, or none of them, belong to the optimal relay set $\sim Rl$; furthermore, if a node does not belong to any optimal betweenness set, italso does not belong to the optimal relay set.

5. DETAILED IMPLEMENTATION OF CAOR

This section introduces the detailed implementation of CAOR, including the initialization phase and the routing phase. The first phase simplifies the network (Algorithm 2), and the second phase computes the minimum expected delivery delay (Algorithm 3) and makes the routing decision (Algorithm 4). Both of the phases are implemented in the distributed way.

```
Algorithm 1 Determine optimal betweenness set
Require: $\{\langle \_v1;l; \_v1;l' \rangle; \cdots ;\langle \_vn;l; \_vn;l'\rangle\}(\_v1;l' \geq \cdots \geq \_vn;l')$
Ensure: $\sim Sl;l' , \langle \_l;l' ;D'l;l' \rangle$
Initialize: $S=\{v1\}$ and $Dl;l' (S)= 1\_v1;l+ 1\_v1;l'$;
for $i=2; \cdots ; n$ do
    $S=S+\{vi\}$;
    Incrementally compute $Dl;l' (S)$ by Eq.(2);
    if $Dl;l' (S)$ increases then
        Break;
return $\sim Sl;l'=S-\{vi\}$ and corresponding $\langle \_l;l' ;D'l;l' \rangle$;
```

5.1 Initialization Phase

The key of the initialization phase is to determine the optimal betweenness sets for each pair of communities. To this end, we introduce a theorem, by which these optimal betweenness sets can be efficiently derived out.
In previous work exploited the fixed point technique to design an efficient algorithm, which can be used to determine the optimal betweenness set $\sim S$. In fact, the results, including a property and the corresponding algorithm, also can be derived from our optimal opportunistic routing rule in Theorem 2.

```
Algorithm 2 CAOR: initialization
Ensure: $G=\langle L;W\rangle$, where $W =\{\langle \_l;l' ;D'l;l' \rangle | l; l'\in L\}$
For each community home $l\in L$ do
    Collect $\_v;l, \_v;l'$ for each $v\in Cl$ and $l'\in L-\{l\}$;
    Use Algorithm 1 to produce $\sim Sl;l'$ and $\langle \_l;l' ;D'l;l' \rangle$;
    Create the virtual link $\xrightarrow{} ll'$ : $\langle \_l;l' ;D'l;l' \rangle$ for each $l'\in L-\{l\}$ and send the link weights to other homes;
```

```
4: Receive the link weights from other homes;
5: Construct the contact graph $G=\langle L;W\rangle$;
```

Now, we present the implementation of the initialization phase by Algorithm 2. Each community home first collects the _ parameters of its community members in Step 1. Then, the home exploits Algorithm 1 to determine the optimal betweenness sets for the message deliveries from itself to other community homes in Step 2. In Step 3, the home produces the virtual links for these deliveries and sends the corresponding weights to other community homes.

5.2 Routing Phase

The routing phase extends the graph, uses the reverse Dijkstra algorithm to compute the minimum expected delays for each home in the extended graph, and then makes the routing decision. The reverse Dijkstra algorithm is shown in Algorithm 3. Steps 1 and 2 are the initialization. In each round, i.e., Steps 4-9, the minimum expected delivery delay of a home is determined.

Algorithm 3 Compute minimum expected delay

```
Require: $G^+ = \langle L^+;W^+\rangle$, $i$, $l_0=d$
Ensure: $D_{i;d}$
1: Set $S=\emptyset$;
2: Let $D_{l_0;l_0}=0$, $S \leftarrow l_0$, and $L^+=L^+-l_0$;
3: for each $l \in L^+$ do
4: Compute $D_{l;l_0}(S)$ according to Eq.(9);
5: Select the smallest one, and let $D_{l;l_0}=D_{l;l_0}(S)$;
6: if $l$ is $i$ then
7: Break;
8: else
9: $S \leftarrow l$, and $L^+=L^+-l$;
10: return $D_{i;d}=D_{l;l_0}$ ;
```

The routing decision of CAOR is shown in Algorithm 4. When a node v visits a community home l, it first construct the extended contact graph of homes G^+ in Step 4 by adding v and d into the graph G, which is generated by home l in the initialization phase. Then, node v uses Algorithm 3 to compute the minimum expected delivery delays $D_{v;d}$ and $D_{l;d}$ in Step 5. The routing decision is made in Steps 6-9.

6. PERFORMANCE EVALUATION

In this section, we conduct extensive simulations to evaluate the performance of the CAOR algorithm using the MSN trace from the Wi-Fi campus network of Dartmouth College [10].

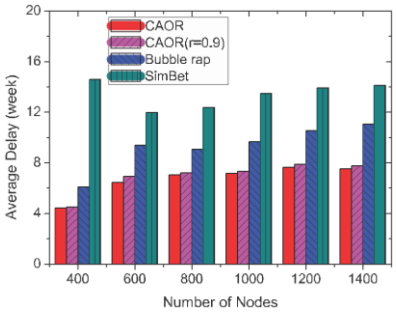
6.1 Simulations on the Real Trace

We adopt the real experimental trace of the Wi-Fi campus network in Dartmouth College [10] in our simulations, since it is one of the most extensive and widely exploited data traces. This trace includes 507 valid APs and uses 6,022 log files to list the records for each node's visit to the APs from 2001-2003. Due to the limit of our PC memory and computation ability, we randomly select partial mobile nodes and Aps from the trace to construct an MSN, while ensuring adequate connectivity of the network. Concretely, the network is built as follows.

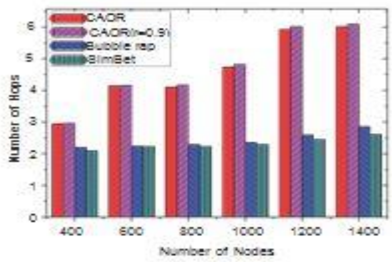
(1) We first derive the number of related nodes for each AP from the trace, and randomly select an AP with enough related nodes as the seed AP, denoted byl_1. Here, a node is said to be related to an AP if there are records about its visit to this AP in the trace.

(2) Secondly, we assume that $l1; \cdots ; li-1$ have been determined, and then we determine the i-th AP. We randomly select an AP from the remaining APs, and compute $|Ci \cap Ck$ for each $k \in [1; i-1]$

(3) We repeat step (2) until $|C1 \cup \cdots \cup Ci| \geq |V|$.If $|C1 \cup \cdots \cup Ci| > |V|$, we randomly remove some nodes to make $|C1 \cup \cdots \cup Ci| = |V|$. Then, we get $L=\{l1; \cdots ; li\}$, and $V=C1 \cdots \cup Ci$. Finally, we compute the _ parameter for each node's visit to each AP, according to the trace.



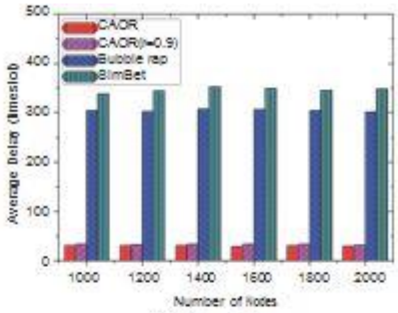
(a) Comparison of average delay



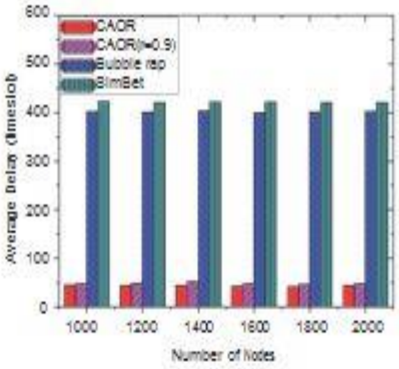
(b) Comparison of average hops

7. RELATED WORK

So far, many traditional DTN routing algorithms have been proposed. These algorithms include flooding-based algorithms (e.g., [6, 7]) and probability-based algorithms (e.g., [8, 9, 10, 11, 12]). Among these algorithms, the MH$*$ algorithm [14] adopts the optimal opportunistic routing strategy, based on global contact information. Compared with this algorithm, the CAOR algorithm adopts the home aware community model and turns the routing problem among mobile nodes into the routing problem among static communities, and therefore, achieves the optimal routing performance only based on community contact information. The maintenance cost of the contact information is far less than the MH$*$ algorithm. This is important because it means that the mobility behaviors of most nodes would not affect the routing performance of the whole network. Moreover, since the network is simplified to be a static network, many previous routing algorithms in static networks, such as wireless sensor networks, can be applied. Social-aware algorithms assume that each node has some social characteristics (such as community, centrality, and similarity, etc.) and then exploits the knowledge to direct the routing decision, so as to improve the delivery ratio. The SimBet [3] algorithm exploits the ego network technique to locally compute the approximate centrality and similarity for each node. It then uses these characteristics to find bridge nodes for the message delivery. The Bubble rap [2] algorithm uses the k-clique algorithm to detect a community, ranks each node by calculating their centrality values, and then exploits the rank values of nodes to direct the routing decision. Besides, the algorithm in [5], a multicasting MSN algorithm, also uses the k-clique technique to detect the communities, and defines the cumulative contact probability of each node as its centrality, based on which, it finds the relay for message delivery.



(a) $|L| = 10$



(c) $|L| = 20$

Fig. 3 Comparisons of delay and hops of CAOR,

8. Conclusion

we model an MSN into some overlap- ping home-aware communities, simplify the routing problem among many mobile nodes into the problem among some static communities, and propose the CAOR algorithm to achieve optimal opportunistic routing. Through theoretical analysis, we find out that optimal opportunistic routing only depends on a few nodes in the network. A change in behavior of most nodes would not affect the routing performance. We can thus achieve the optimal routing performance at a very low maintenance cost. Compared with previous social-aware algorithms, the optimal and predictable routing performance is the biggest advantage of the CAOR algorithm.

References

[1]. Wu, M. Xiao, and L. Huang, “Homing spread: Community home-based multi-copy routing in mobile social networks,” in *IEEE INFOCOM*,2013.
[2]. P. Hui, J. Crowcroft, and E. Yoneki, “Bubble rap: social-based forwarding in delay tolerant networks,” in *ACM MobiHoc*, 2008.
[3]. E. Daly and M. Haahr, “Social network anal- ysis for routing in disconnected delay-tolerant manets,” in *ACM MobiHoc*, 2007.
[4]. M. Ibrahim, P. Nain, and I. Carreras,Analysis of relaprotofor throwbox-equipped dtns in WiOPT, 2009.
[5]. D. Kotz, T. Henderson, I. Abyzov, and J. Yeo, “CRAWDAD data set dartmouth / campus (v. 2009-09-09),” Sep. 2009.

[6]. J. Wu and Y. Wang, “Social feature-based multipath routing in delay tolerant networks,” in IEEE INFOCOM, 2012.

[7]. W. Gao, Q. Li, B. Zhao, and G. Cao,“Multicasting in delay tolerant networks: A social network perspective,” in ACM MobiHoc, 2009.

[8]. Q. Li, S. Zhu, and G. Cao, “Routing in social selfish delay tolerant networks,” in IEEE INFOCOM,2010.

[9]. R. Lu, X. Lin, and X. S. Shen, “Spring: A socialbased privacy-preserving packet forwarding protocol for vehicular delay tolerant networks,” in IEEE INFOCOM, 2010.

[10]Vahdate and D. Becker, “Epidemic routing for partially-connected ad hoc networks,” Duke University, Tech. Rep. CS-2000-06, June 2000.

[11]T. Spyropoulos, K. Psounis, and C.Raghavendra,“Spray and wait: An efficient routing scheme for intermittentlyconnected mobile networks,” 2005.

[12].Lindgren, A. Doria, and O. Schelen, “Probabilistic routing in intermittently connected networks,”Lecture Notes in Computer Science, vol.3126, no. 1, pp. 239–254, 2000.