Spatial and Temporal Communities-based Routing Algorithm in Human Social Network

Tao ZHOU¹, Hongbing XU¹, Yuanyuan CHEN²

¹University of Electronic Science and Technology of China, Chengdu 611731, China
²Sichuan University, Chengdu 610064, China

Abstract

Human Contact Network is a kind of DTN (Delay Tolerant Networks), and is closely related with people’s activities. Owing to its intermittent connectivity and unique human mobility patterns, how to transmit data in an effective way is a challenging problem for DTN. In this paper, considering the regularity of people activities, we introduce a THM (Trip History Model) to record the activities of people. Through the spatial and temporal characteristic of people stored in THM, combined the community theory, the further moving destination can be predicted. In order to transfer data more reasonable, Spatial and Temporal Communities-based Routing Algorithm (STCRA) has been proposed, which data will be transferred to superiority nodes through the community history information. A message management mechanism is proposed to effectively control the number of messages and reduce the network overhead. Our simulation results show that, compared with Flooding and LABEL algorithms, STCRA effectively improves the routing performance of the network.

Keywords: Human Contact Network; Delay Tolerant Network; Routing; Human Social Network

1 Introduction

An opportunistic network are considered as Delay Tolerant Network (DTN) [1], where communication opportunities (contacts) are intermittent, so an end-to-end path between the source and the destination may never exist. A store-and-forward mechanism is proposed and commonly used to transfer data in DTNs. Pocket Switched Network (PSN) [2] is a kind of DTNs, which is composed of mobile communication devices and focus on coping with human mobility patterns. With the rapid increase in using mobile devices, PSN is likely to be a popular communication network in the near future.

Many routing schemes have been proposed for PSNs. As the dynamic topology and intermittent connectivity of PSNs, sensor nodes can only communicate by using store-ferry message switching when they meet. The most simple and basic scheme is Direct Transmission, which the mobile
nodes only transfer data when they meet the sink node. Thus, it has a low delivery ratio and a long transmission delay.

Vahat and Becker [3] propose epidemic routing protocol, which is the representative flooding routing protocol. It attempts to give all contacted nodes a copy to every message through the communication with them to achieve high delivery ratios. Flooding protocol does not require any prior global or local knowledge about the sensor network, but is also has some defects. If the buffer of each node is large enough, flooding protocol can get a higher delivery ratio and a lower transmission delay, but it will consume more energy to send and receive message copies and occupy a large proportion of network bandwidth.

In recent years, social-based approaches [4], which attempt to exploit social behaviors of PSN nodes to make better routing decision, have drawn tremendous interests in PSNs routing design. Community is an important concept in ecology and sociology [5, 6, 7], which studies the interactions between species/people at spatial and temporal scales. Community is a group of interacting people living in a common location, and reflects social relationship among people.

Hui and Crowcroft [8] propose a LABEL routing algorithm, which assumes every node belongs to a community and marks nodes in the same community with the same label. When a node needs to send a message, it will only select the nodes within the communication range which belong to the destination node community as the next hop node. However, when the source node is far from the destination node community, the message cannot be transferred. For this reason, literature [8] presents a BUBBLE routing scheme, which takes the node activities into account. Each node has global level and a local level; the higher its global level is, the more active and the node will be and thus having a stronger ability to transfer data.

Daly [9] presents a delay tolerant network routing technique (Simbet) based on the small world social-phenomenon, which comes from the observation that individuals are often linked by a short chain of acquaintances. According to the betweenness and the social similarity value to jointly decide the Simbet value of each node. When two nodes meet in the network, the message will be transferred to the node with larger Simbet value in order to improve the delivery performance. Simulations using three real trace data sets demonstrate that the achieved delivery performance may be closed to that of Epidemic Routing but the overhead can be significantly reduced.

Because of the dynamic topology and the intermittent connectivity of PSN, people can only communicate by using store-and-forward message switching when they meet. In order to get a better routing algorithm of PSN, some scholars have studied in human mobility patterns based on human historical records. Crowcroft [10] shows that many people will move according to some certain rules. For instance, they are likely to move on some certain paths periodically. In addition, a project named Mobidrive [11] records the movement of people in six weeks and found that people’s movement between 4 to 6 major locations may occupy more than 70% of all the movements.

Based on the research results of human social systems, Trip History Model (THM) [12] is proposed to record people’s actual trajectory. Through the spatial and temporal information stored in THM, the places to which the person often goes can be gotten, and the communities to which the person belongs can be analyzed. We use these empirical data to construct the history temporal communities at different time periods. The history spatial and temporal communities will help calculate the probability the node communicate with the sink node, and improve the delivery ratio.
The rest of the paper is organized as follows. Section 2 introduces the network model. Section 3 introduces the proposed routing scheme. Section 4 introduces Mechanism of Message Management. Section 5 presents the simulation results and discussion. Finally, Section 6 concludes the paper.

2 Network Model

This paper assumes that the senor nodes are distributed in a two-dimensional region A, as Fig. 1. There are several streets in the region A. Some of streets are from south to north and the other streets are from east to west. The distance between two streets is random. All sensor nodes which have a unique ID are isomorphic and the moving trajectories of them are along the streets. Some people activity sites are randomly deployed in the streets, such as offices, restaurants, home, schools and so on. Some static sink nodes are randomly distributed in the region. For simplicity, all the sensor nodes and sink nodes have the same transmission radius R.

To get a closer simulation scenario with people’s actual activities trajectory the movements of all the sensor nodes follow the Trip History Model (THM) [12]. THM records the moving trajectory of people. In addition, each sensor node has a home location, and the sensor node will often move from home in the morning and reach home at night.

Fig. 1: Simulation scenario and node mobility model

3 The Proposed Routing Scheme

We consider people’s activities have certain rules and can be divided into different community according to spatial and temporal correlations. For instance, before a person leave home, he is in a family community, and he is in a colleague community when he reaches his office. In order to clearly describe the community which the person belongs to at different time, we divide a day into a discrete time set, that is Time $(< 8, 8 − 9, 9 − 11, 11 − 12, 12 − 13, 13 − 15, 15 − 17, 17 − 19, > 19)$. As Fig. 2, in different time set T, the sensor network can be divided into some communities. These communities information can help transfer messages and improve the delivery ratio.
If any two node $V_i$ and $V_j$, if the distance between them is less than $R$ and the duration that they may communicate with each other is longer than $\theta$, they are called neighbors, and let neighbor $(V_i)$ stand for the neighbor of $V_i$. The community including $V_i$ and $V_j$ is $Community(c) = \{V_i, V_j\}$. Meanwhile, we arbitrarily select a sensor node from $V_i$ and $V_j$ as the community head which maintains the routing information of the whole community. The community structure can be described as an undirected graph $G=(V,L)$ saved by the community head, which $V$ is the set of all nodes in the network and $L$ is the set of connections between nodes. When any nodes enter or leave the community, a special message will send to the community head and the community head will change $G=(V,L)$ according to the special message. If a sensor node $V_k$ which does not belong to another community becomes the neighbor of $V_i$ or $V_j$, $V_k$ will be added to the community, that is $Community(c) = \{V_i, V_j, V_k\}$. If $V_k$ belongs to another community, combine these two communities into a community, and let one community head send the community network structure graph $G=(V,L)$ to another community head.

When a sink node becomes a member of a community, we call the community as the valid community. By the trip history of people, we can record the history valid community to which a mobile sensor node $V_i$ belongs as shown in Table 1. Frequency is the times of $V_i$ becoming the member of the community.

<table>
<thead>
<tr>
<th>No.</th>
<th>Time</th>
<th>Valid community</th>
<th>Sink node in community</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8-9</td>
<td>Community(1)</td>
<td>S1</td>
<td>15</td>
</tr>
<tr>
<td>2</td>
<td>8-9</td>
<td>Community(3)</td>
<td>S3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>9-10</td>
<td>Community(4)</td>
<td>S4</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>17-19</td>
<td>Community(2)</td>
<td>S2</td>
<td>12</td>
</tr>
</tbody>
</table>

From the trip history records above, we can get the community in which $V_i$ is often at different time set. $V_i$ is the member of community (1) in time set [8-9] for 15 times, and is the member of community (3) for 3 times. Let $P_T(v_i)$ stand for the probability that $V_i$ sends messages to sink node at time set $T$ through the community and $P_T(v_i)$ can be calculated as:

$$P_T(v_i) = \sum F_T(V_i) \cdot (1 - \delta)^{\tau_T}/S_T$$
Where $F_T(v_i)$ is the times of $V_i$ becoming the member of the valid community at time set $T$, and $S_T$ is the total statistics days. $\delta \in [0,1]$ is a attenuation coefficient which controls the impact of valid community number on the $P_T$. $\tau T$ is the number of valid community in time set $T$. If there exist some valid community in the same time set instead of one and the frequency of them may be little, which means people’s activities in this time set is not very regular, then probability that messages sent to destination should be reduced.

Assume node $V_i$ meets node $V_j$ in time set $T$, and both nodes have messages to be transferred. Two nodes inquire the probability sending messages to sink node at time set $T$ and $T+1$ of others through a simple handshake message. If $P_T(V_i)$ or $P_{T+1}(V_i)$ is larger than $P_T(V_j)$ or $P_{T+1}(V_j)$, $V_j$ will send messages to $V_i$ and vice versa. We compare the probability of messages which is sent to the sink node at current time set and the next time set of $V_i$ and $V_j$, and then the messages will be transferred to the sensor node which has a greater possibility to communicate in near future with sink node.

4 Mechanism of Message Management

Each sensor node has a message storage queue which stores messages to be sent. Taking the data redundancy in sensor networks into account, assume each message contains a storage TTL (Time-To-Live) domain $\rho$. The initial value of $\rho$ is 0, and $\rho$ will increase as time passed by. Assume the network delay tolerance threshold (of the network delay tolerance) is $\eta$. If a message’s TTL $\rho$ is more than network delay tolerance threshold $\eta$, the message will be immediately discarded. In this way, it will become possible to avoid retaining too many messages in the sensor network, which may bring about the waste of network bandwidth and consumption of much energy of sensor nodes.

The message storage queue has three sources shown in the Fig. 3. 1) Sensor nodes collect message generated by itself. 2) Sensor nodes receive the messages from other sensor nodes. 3) Sensor nodes send messages to other sensor nodes, and save the copies of these sent messages in the storage queue. Messages are sorted by the survival time from short to long in the message storage queue. The message which has short survival time will be put into the front (head) of the queue and has high priority to be sent.

![Fig. 3: Source of messages in the queue](image-url)
5 Simulation

Extensive simulation has been carried out to evaluate the performance of STCRA. We compare the performance of STCRA with that of Flooding and LABEL algorithms.

5.1 Simulation setting

Assume that the simulation area is $2000 \times 2000$, and 200 mobile sensor nodes and 20 sink nodes are randomly deployed in it. The network bandwidth is 8 Kbps. The default message queue is 300, and the survival time of the message is 200 seconds. In addition, there are 60 sites including 20 "homes", 15 schools, 15 offices, and 10 shops randomly distributed in the simulation area. We define three types of nodes, named students, workers, and other personnel. For workers, the first everyday activity is going to office as it is going to school for students. Other personnel’s first everyday activity is going to other sites except for schools and offices. The first activity is randomly selected between 6:00 AM to 8:00 AM. All the people will go home and rest before 12:00 PM. The simulation duration is 90 days, the former 60 days are used to create trip history model and the remaining 30 days are used to simulate. The simulation results are as follows.

5.2 Performance comparison

The network topology is closely associated with the sensor node density. We change node density by increasing the total number of sensor node, and observe the impact on the performance of these algorithms. The total number of nodes is increased from 100 to 300, and the experiment results are shown in Fig. 4. As shown in Fig. 4(a), with more sensor node deployed in sensor network, delivery ratio of these three algorithms will increase. Due to an increase of node density, more sensor nodes can help transfer messages to the sink node, and so delivery ratio will increase. Flooding has a higher delivery ratio than other algorithms when there is little number of nodes existing in the network. Because of low node density, nodes have enough buffers to save the message copies and more messages can be transferred to sink node. However, when more nodes existing in the network, amount of message copies will be created and some of them will be discarded, then the delivery ratio of Flooding algorithm is not ideal at this stage. Fig. 4(b) shows the node density’s impact on the average copies of each message. Average copies will increase with the growth of node density. More message copies will be transferred to neighbor nodes and average copies will increase. The average copies of each message in LABEL are the least, because the mobile nodes only transfer messages to the other node which is belonging to the same group with it. In Fig. 4(c), the average delay in the three algorithms will decrease with the increase of sensor node number. The average delay of STCRA is always less than that of the other algorithms. When the node density increases, there are more message copies created in the sensor network, and more chance to be transferred to the sink node with a short delay.

The experiment below studies the performance of the sensor network with different message queue size. We change message storing queue size from 50 to 500, and the simulation results are show in Fig. 5. Message storing queue size means the maximum number of messages which can be carried by the mobile sensor node. As shown in Fig. 5(a), with the increase of queue size, the delivery ratio in these three algorithms will increase. When the mobile sensor node has a bigger queue size, more messages can be carried by the node, and messages have more chance to be
transferred to the sink node. Fig. 5(b) shows the queue size makes little impact on LABEL and STCRA and makes great impact on Flooding. There is little message copies existing in STCRA, even in a large queue size. Fig. 5(c) shows the average delay of STCRA is less than that of the other algorithms. When the queue size increases, messages can be stored in the sensor nodes and survive longer until it is discarded. Hence, the average delay will increase a little.

6 Conclusions

In this paper, we have studies the communication in the human contacted networks, which is a combination between Delay Tolerant Networks and people society network. The trip history model has been introduced to record the activities of people who are treated as mobile sensor nodes. By means of the observation in people’s mobility characteristics, the community which the person often goes can be gotten and used to predict the further destination. In order to describe people’s activities more detail, we proposed spatial and temporal communities to record the place and the communities mostly to go at different time. Through the valid spatial and temporal communities selecting, messages will be transferred to sink node more easily, and then will improve the delivery ratio. A message management mechanism is proposed to effectively control the number of messages and reduce the network overhead. Combined with judgments
of human’s trajectory, STCRA effectively avoid the blindness in data forwarding. According to simulation results, STCRA has a better performance, which can obtain a higher delivery ratio with lower energy consumption and transmission delay.

Acknowledgement

This work is supported by the National Natural Science Foundation of China under Grants nos. 60903158, 61170256, 61173172, 61103226, 61103227.

References