

## Research Article

# Spatial and Temporal Correlations-Based Routing Algorithm in Intermittent Connectivity Human Social Network

Zhou Tao,<sup>1</sup> Xu Hongbin,<sup>1</sup> Liu Ming,<sup>2</sup> Liu Nianbo,<sup>2</sup> and Gong Haigang<sup>2</sup>

<sup>1</sup> School of Automation, University of Electronic Science and Technology of China, Chengdu 611731, China

<sup>2</sup> School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

Correspondence should be addressed to Liu Ming, wing.lm@gmail.com

Received 13 December 2011; Accepted 14 January 2012

Academic Editor: Mo Li

Copyright © 2012 Zhou Tao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The social network formed by people is one of the key applications of Delay-Tolerant Network (DTN). Owing to its intermittent connectivity and unique human mobility patterns, how to transmit data in an effective way is a challenging problem for the social network. In this paper, we propose the idea of Trip History Model (THM) which establishes a model on a single person's mobility, and then a Spatial and Temporal Correlations-Based Routing Algorithm (STC) is proposed. In STC, the node delivery probability is calculated according to both a node's current moving prediction and its history record to give guidance for message transmission. Our simulation results show that, compared with LABEL and PROPHET algorithms, STC effectively improves the routing performance of the network.

## 1. Introduction

The application of wireless network adapters in portable communication devices, such as mobile telephone, personal digital assistants, and laptops, can provide more chances to access global information resources. These portable devices can be connected to dynamic networks. By such networks, people can transfer information to each other. The movement of people and the limited communication range make the network unable to always guarantee its own connectivity. To make communication possible, DTNs (delay tolerant networks) [1] were proposed, which overcome the problems associated with intermittent connectivity, long or variable delay, asymmetric data rates, and high error rates by using store-and-forward message switching. PSN (pocket switched network) is a kind of DTNs, which is composed of mobile communication devices and has broad application in the future.

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. Currently, many proposed PSN routing algorithms are based on historical records, and data is transmitted to the node which can constantly communicate with destination node. Since

these algorithms fail to predict the trajectories of mobile nodes, they cannot guarantee the reasonableness of data forwarding. In addition, [2] shows that many people will move according to some certain rules. For instance, they like moving to specific sites periodically; in this case, regular space and time rules can help to guide routing in PSN.

Based on human movement discipline, this paper proposes Trip History Model (THM) to automatically predict the destination of people by studying their trip history recorded by portable communication devices. Meanwhile, according to the spatial and temporal characteristics of human movement, STC (Spatial and Temporal Correlations Routing Algorithm) is designed to forward the message to next node according to the prediction of the current node. Simulation results show that STC effectively improves the routing performance.

## 2. Related Work

In recent years, there has been extensive research on routing protocols in delay-tolerant networks. Vahdat and Becker [3] present epidemic routing protocol, which is the representative flooding routing protocol. It attempts to give all nodes a copy of every message through the communication

with them to achieve high delivery ratios. Flooding routing protocol requires no knowledge about the sensor network, but it also has some defects. For instance, since there exists huge amounts of copies of messages, a large proportion of network bandwidth and node buffer resources will be occupied. Meanwhile, more energy of the sensor nodes will have to be consumed to send and receive these copies messages. Based on the research on human society, some scholars proposed the community-based intermittent connectivity routing algorithm.

Literature [4] proposes a Replication-based Efficient Data Delivery (RED) scheme, which consists of two key components for data transmission and message management, respectively. Data transmission scheme decides when and where to transmit data messages based on the delivery probability, and message management scheme makes decisions on the optimal erasure coding parameters according to its current delivery probability. However, the calculation of erasure coding parameters is not a refined one, and a lot of data blocks split by RED will consume lots of energy of the sensor nodes.

Literature [5] presents a LABEL routing scheme, 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, Hui and Crowcroft [5] propose a BUBBLE routing scheme, which takes the node activity into account. Each node has a global level and a local level; the higher its global level is, the more active the node will be and thus having the stronger ability to transfer messages.

In addition, Daly and Haahr [6] present 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. The betweenness centrality and the social similarity value are 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.

### 3. Analysis

In recent years, great progress has been made in the research of human complex social systems. Based on the tracking data of the wireless mobile network users, Hsu et al. [7] sum up two characteristics of people's movement, that is, preferring specific locations and going to the same location periodically. Based on the tracking data of thousands of anonymous mobile phone users, González et al. [8] propose people's activities have regular space and time rules. These research results prove the movement of people is nonrandom and it has regular space and time rules. In addition, a project named

Mobidrive [9] recorded the movement of people in six weeks and found that human movement between 4 to 6 major locations may occupy more than 70% of all the movements. Therefore, the movement of people can be described as a spatial one between regular locations and thus the conclusion that the movement of many persons can be predicted.

The current human movement models fall into the following categories: firstly, the random movement model, including random waypoint, random walk, and so forth. In the random movement model, node randomly select the destination, direction, and thinking times, and so forth. The movements of all nodes in the models mentioned before follow the same rule. These models have simple structure and can be clearly defined, but they cannot reflect the movement features of people in real world for its lack of authenticity. Second, modeling and analyzing the movement features of people are done based on the collection of the real trajectory. Chaintreau [10] considers intercontact time of portable wireless devices nearly obeys the power-law distribution rather than the original exponential distribution. For such models, the main difficulty lies not in the model itself, but in how to organize data and obtain the implied movement mode through a series of mathematical methods. Third, the movement model is established based on community. Literature [11] divides people into different groups based on the social relations between them. These groups are randomly placed in the simulation area and the nodes select their destination and move according to certain rules. Literature [12] introduces agenda-based routing protocol (ARP), which considers the relationship between human social activities and movement locations based on the survey data of US Transportation Department, but it does not consider the two characteristics of human movement, that is, spatial and temporal. The models above have their defects, and therefore we propose Trip History Model that takes into account the two characteristics of human movement to predict the human activities, which can be easily saved and processed.

## 4. Network Model

*4.1. Trip History Model.* This paper assumes that the sensor nodes are distributed in a two-dimensional  $M \times M$  square region  $A$ . All the sensor nodes have the same transmission radius  $R$ . Network can be described as an undirected graph  $G = (V, L)$ ;  $V$  is the set of all nodes in the network, and  $L$  is the set of connections in the network. For any two nodes  $V_i$  and  $V_j$ , if the distance between them is less than  $R$ , there exists a connection between them and let *neighbor* ( $v_i$ ) stand for the neighbor of  $V_i$ .

In addition, the movements of all the sensor nodes follow the Trip History Model. The power of portable devices can be replenished in time, so we assume that the energy of sensor nodes is infinite.

We consider the human activities have certain rules and can be described as a series of data including time and location. All the records of time and location of a person's activities constitute his travel history. Since a person's activities include some repetition every day, we save a person's trip

in a day as a record cycle. Meanwhile, a day can be divided into two discrete sets, that is, Day (Mon., Tues., Wed., Thurs., Fri., Sat., Sun.) and Time (<8, 8-9, 9-11, 11-12, 12-13, 13-15, 15-17, 17-19, >19). For instance, the trip history of a person can be described as follows:

- (1) Mon. 7:40 move from home to workplace along a certain way.
- (2) Mon. 17:30 go home from the office and pass by the supermarket for some shopping.
- (3) Sat. 10:20 move from home to the shopping mall.
- (4) Sat. 18:30 go home from the shopping mall and pass by the restaurant to eat something.

The fourth trip record presents a stopover, and such behavior is true in real life. The trip history above can be described as in Table 1.

In addition, for getting easy computing and saving, we adopt some methods to change continuous location into discrete value. Home or Office, or Supermarket, and so forth is not a real GPS location, but a neighborhood GPS location set. People probably play and move near home and office, and the value of GPS may be some measurement errors, so we can consider the positions within a certain range (100 meters) as the same position.

When a person starts moving, the portable device records the current time and position until he reaches the destination. After a trip is finished, the portable device compares the trip with the shortest way according to the map. If the trip is the shortest way, the portable device will save the time, source, and destination and delete other trip history data. Otherwise, it records the time, source, destination, and the midpoints which the person passed by.

**4.2. Decision Tree.** We use a decision tree to save Trip History Model and help make some decisions on transferring messages in order to improve the delivery performance. As Figure 1 shows, THM can be presented as the branches and leaves of the decision tree.  $P_{kq}$  is the probability that  $q$  is destination and can be calculated as

$$P_{kq} = \frac{f_{kq}}{N_k}, \quad (1)$$

where  $f_{kq}$  is the number of trips according with specific time, midway point, and destination choices at leaf node  $kq$ . Since category  $kq$  gives a final destination choice,  $P_{kq}$  represents the probability of moving toward the destination in history. Since the car knows current Source, Day, and Time values, it can find a Destination choice with a maximum  $P_{kq}$  value in the tree as an initiate prediction.

During the movement of people, the portable device will continue to check the current position and identify, if people are still on the way, the predicted midpoint, or destination. If the position disagrees with the predicted route, the new destination prediction will be calculated as

$$P_{kq} = \begin{cases} 0 & \text{if } q \text{ is infeasible,} \\ \frac{f_{kq}}{\sum f_k} & \text{other,} \end{cases} \quad (2)$$

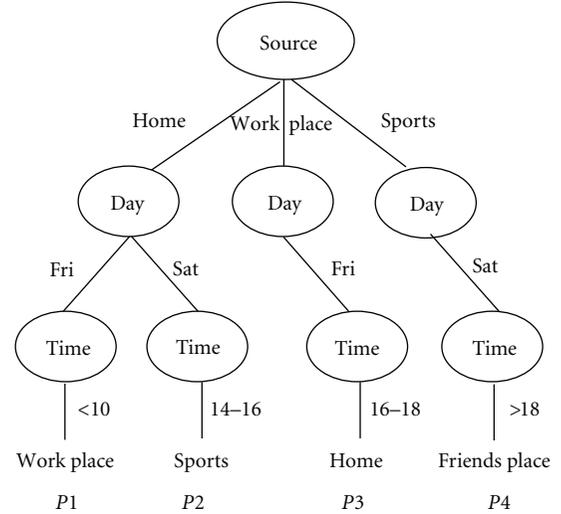


FIGURE 1: The structure of decision tree with trip history model data.

where  $\sum f_k$  is the total number of the rest of the trips after removing the infeasible ones.

## 5. Spatial and Temporal Correlations Routing Algorithm (STC)

The trip history data is collected by the Trip History Model, including spatial and temporal information. To make use of the history data and to predict the movement of the people (nodes) in near future, we introduce spatial and temporal correlations routing algorithm (STC).

**5.1. The Encounter Information Table.** Assume each node  $n$  saves the encounter information including ID, encounter frequency, and predicted destination, as in Table 2.

Assume the node  $V_i$ 's predicted destination is changed, so  $V_i$  will send a special message to all its neighbor nodes. This message includes the ID of the node  $V_i$  and the new predicted destination. And the message disseminates as epidemic-like approaches, until all nodes in the network receive this message. When node  $V_j$  receives this message, it will do the following.

- (1) If  $V_i$ 's information has been included in  $V_j$ ,  $V_j$  updates  $V_i$ 's predicted destination.
- (2) If no,  $V_j$  adds  $V_i$ 's information in its encounter information table, and sets the Encounter Frequency to zero.

**5.2. Transmission Probability.** Let  $P_{io}$  stand for the transmission probability of node  $V_i$  transferring the message to node  $V_o$ :

$$P_{io} = \frac{\theta}{180^\circ}, \quad (3)$$

TABLE 1: Records of the trip history.

No.	Day	Time	Source	Midpoint	Destination
1	Mon.	<8	Home		Office
2	Mon.	17–19	Office	Supermarket	Home
3	Sat.	9–11	Home		Shopping Mall
4	Sat.	17–19	Shopping Mall	Restaurant	Home

TABLE 2: Encounter information table.

ID of encounter node	Encounter frequency	Predicted destination
5	2	D1
21	5	D2
3	3	D3
12	1	D1

where  $\theta$  is the angle between two tangents of the predicted destination of node  $V_i$  to the transmission radius of predicted destination of node  $V_o$  as Figure 2. According to trigonometric formulas, we can get

$$\theta = 2 \arcsin\left(\frac{R}{d}\right), \quad (4)$$

where  $R$  is the transmission radius of node  $V_o$ , and  $d$  is the distance between the predicted destination of node  $V_i$  and the predicted destination of node  $V_o$ .

**5.3. Message Transmission.** Suppose node  $V_i$  needs to send message to node  $V_o$ . Firstly,  $V_i$  would know  $V_o$ 's predicted destination. If there is not the  $V_o$ 's information in  $V_i$ 's encounter information table,  $V_i$  sends inquiring message to its neighbor nodes. If its neighbor nodes also have not the  $V_o$ 's information, it will inquire to more nodes, until it gets the  $V_o$ 's information. Second,  $V_i$  gets its neighbor's predicted destination by shake hands; if its neighbor nodes' predicted destination is same as  $V_o$ , it will send the message to these neighbor nodes, if not,  $V_i$  will send the message to the neighbor node which has the higher transmission probability of  $V_o$  than  $V_i$ .

## 6. Simulation Results

**6.1. THM Performance.** First, the performance of trip history model is evaluated through the investigation of human trip. In order to get human real trip data, we conducted a 9-week investigation. Invite 5 campus staff members to participate in the survey, and these volunteers have smart handheld devices. In the survey, the drivers need to record the start time, start position, stop position, and midpoint position. The collected trip history data is shown in Table 3. We input these data into the electric map of Chengdu and establish trip history model for each volunteer. Finally, we divide the trip history in Table 3 into two parts. The trip history data in 1–8 weeks regarded as the training set is used to generate the decision tree, and the trip history data in the 9th week

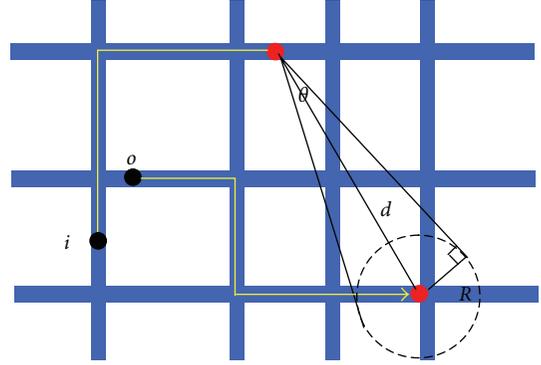


FIGURE 2: Transmission probability calculation (blue lines present the streets).

TABLE 3: Trip history collection of investigation.

People	1–8 week	9 week	Total	Average
A	155	17	172	2.73
B	214	29	243	3.86
C	169	18	187	2.97
D	181	25	206	3.27

considered as the target set is used to evaluate the predicted result.

When the vehicles began to travel, completed a quarter of the trip, half of the trip, three quarters of the trip, and all the trip, the hitting rate compared the predicted destinations of the decision tree with the real destinations, and the results are shown in Figure 3. The average starting accuracy rate is 65.83%, and with the movement of vehicles, the average prediction accuracy rate increases gradually, from 77.56% to 82.77%, to 91.42%, and finally to 97.04%. Although individuals' daily activities are highly regular, the trip in 1–8 weeks cannot fully cover the schedule in the 9th week. Unpredicted circumstances cannot be avoided, so the final average prediction accuracy cannot reach 100%.

**6.2. STC Performance.** To compare STC to other algorithms, we simulate STC, LABEL, and PROPHET algorithms. In PROPHET algorithm, each node calculates the transmission probability for different destination nodes and only transmits the message to the node which has larger transmission probability for the destination nodes.

**6.2.1. Simulation Setting.** Assume that the simulation area is  $2000 \times 2000$ , including 200 nodes and 60 sites which

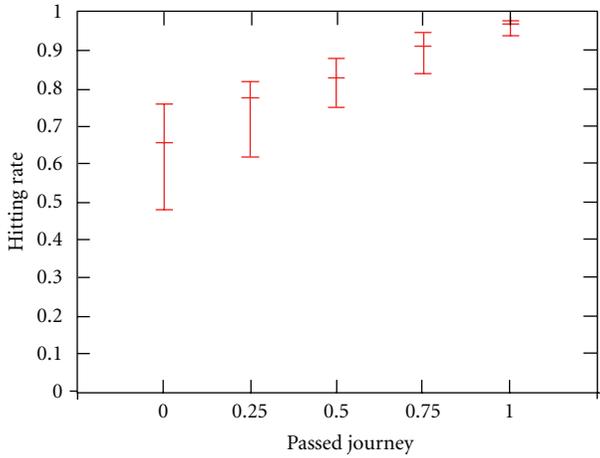


FIGURE 3: The prediction accuracy of trip history model.

TABLE 4: The performance comparison under default parameters.

	STC	LABEL	PROPHET
Delivery ratio (%)	85.6	48	83
Average copies for each message	7.3	1.5	10.5
Average delay(s)	108.7	131.4	110.2

consist of 20 “homes”, 15 schools, 15 offices, and 10 shops or restaurants. These sites are randomly distributed in the simulation area. In addition, we define three types of nodes, namely, students, workers, and other personnel. For students, the first everyday activity is going to school as it is going to office for workers. Other personnel’s first everyday activity is going to other sites except 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. Assume that the network bandwidth is 8 Kbps. The message queue length of the nodes is 500, and the survival time of the message is 300 seconds. The simulation duration is 3 days, and the simulation results are as follows.

**6.2.2. The Position of Nodes at Different Time.** Figures 4 and 5 show the nodes’ positions in the simulation area at 8:00 AM and 12:00 AM. The positions of nodes in Figure 4 are different from those in Figure 5, and this shows nodes have already moved. The positions of the nodes are randomly distributed at different times because the sites are randomly distributed in the simulation. In addition, we also find that the density of nodes in the central region is bigger than that in the edge region, and the nodes tend to move to the central region. Considering the actual situation, there will be an uneven distribution of nodes, for example, more nodes should be at “home” in the evening.

**6.2.3. Performance Comparison.** The following is the performance comparison of three algorithms under default parameters in Table 4. STC’s delivery ratio is bigger than PROPHET’s delivery ratio, but the average number of the copies of each message in STC is lower than that in

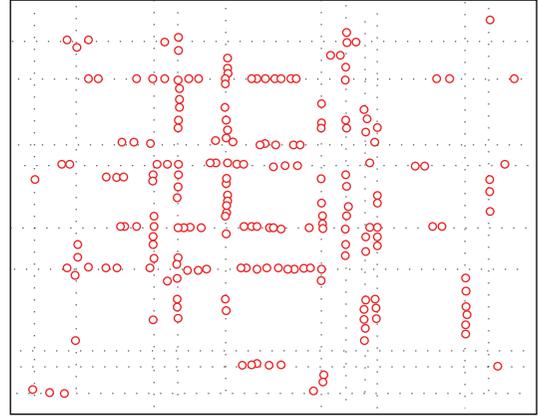


FIGURE 4: Node distributions at 8 AM.

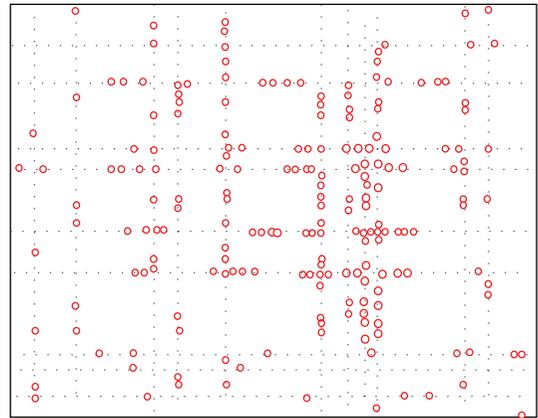


FIGURE 5: Node distributions at 12 AM.

PROPHET, and this shows the former’s average transmission energy consumption is smaller. In addition, we do not consider the energy consumption of exchanging routing tables between two nodes in PROPHET. If the network has more nodes, the energy consumption will be enormous. In STC, when two nodes meet, they do not need to exchange the routing information and much energy will be saved. Table 4 shows the average delivery ratio of LABEL is the lowest and the average delay of LABEL is larger than that of STC, because nodes only transmit collected data to the node with same predicted destination with them in LABEL, and so message transmitted slowly and transmission delay will be longer.

**6.2.4. TTL’s Impact on Performance.** We change TTL (Time To Live) of messages from 50 s to 300 s, and the simulation results are shown in Figures 6(a)–6(c). In the figures, with the increase of TTL, the average delivery ratio of these three algorithms will increase. There is little difference between STC and PROPHET in terms of the delivery ratio, but compared with LABEL’s delivery ratio, theirs are much higher. It shows that the longer survival time the messages have, the more opportunities they get to be transferred to the destination node. Figure 6(b) shows the average number of the copies of each message will increase with the growth

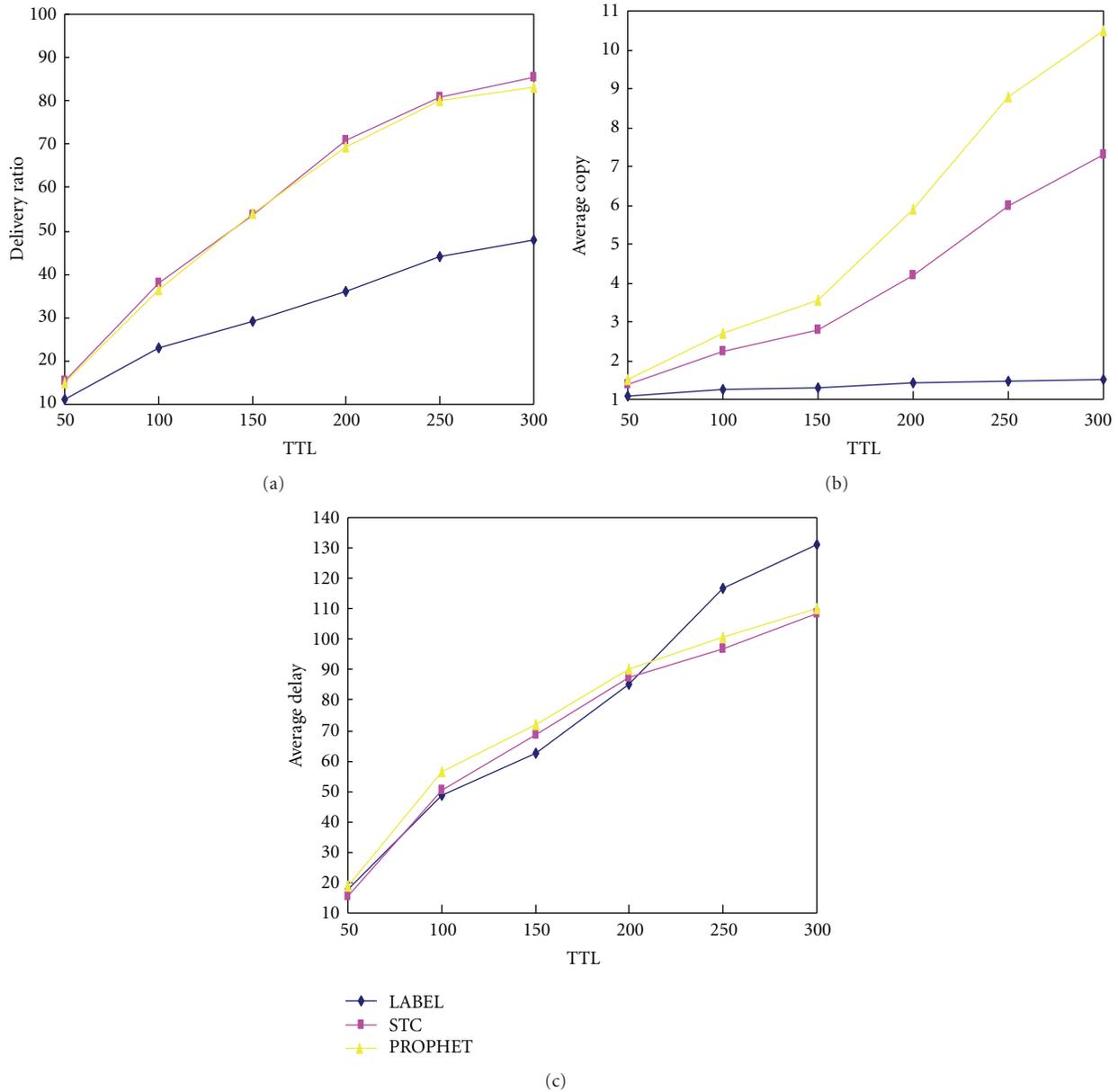


FIGURE 6: TTL's impact on performance. (a) Average delivery ratio. (b) Average copies number. (c) Average delay.

of TTL. Although there is little difference between STC and PROPHET in terms of the delivery ratio, the average number of the copies of each message in STC is much lower than that in PROPHET, which implies that STC has less energy consumption. Figure 6(c) shows that average delay increases with the growth of TTL, and the average delay in STC is slightly lower than that in PROPHET.

**6.2.5. Node Density's Impact on Performance.** The total number of nodes is increased from 100 to 300 to change the node density, and the simulation result is shown in Figures 7(a)–7(c).

Figure 7(a) shows that in all the three algorithms the delivery ratio will increase with the growth of node density.

The increase in the number of nodes will cause a change in the transmission range of each node; as a result, there will be more next hop nodes satisfying the transmission conditions, and thus increasing the delivery ratio. Figure 7(b) shows the average number of the copies of each message will increase with the growth of node density. Although there is little difference between the delivery ratio in STC and that in PROPHET, the average number of the copies of each message in STC is much smaller than that in PROPHET. Figure 7(c) shows the average delay in STC is much lower than LABEL with the number of nodes increasing. With the increase in the number of nodes, there will be more message copies in the network, and it will reduce the number of hops between the source node and the destination node.

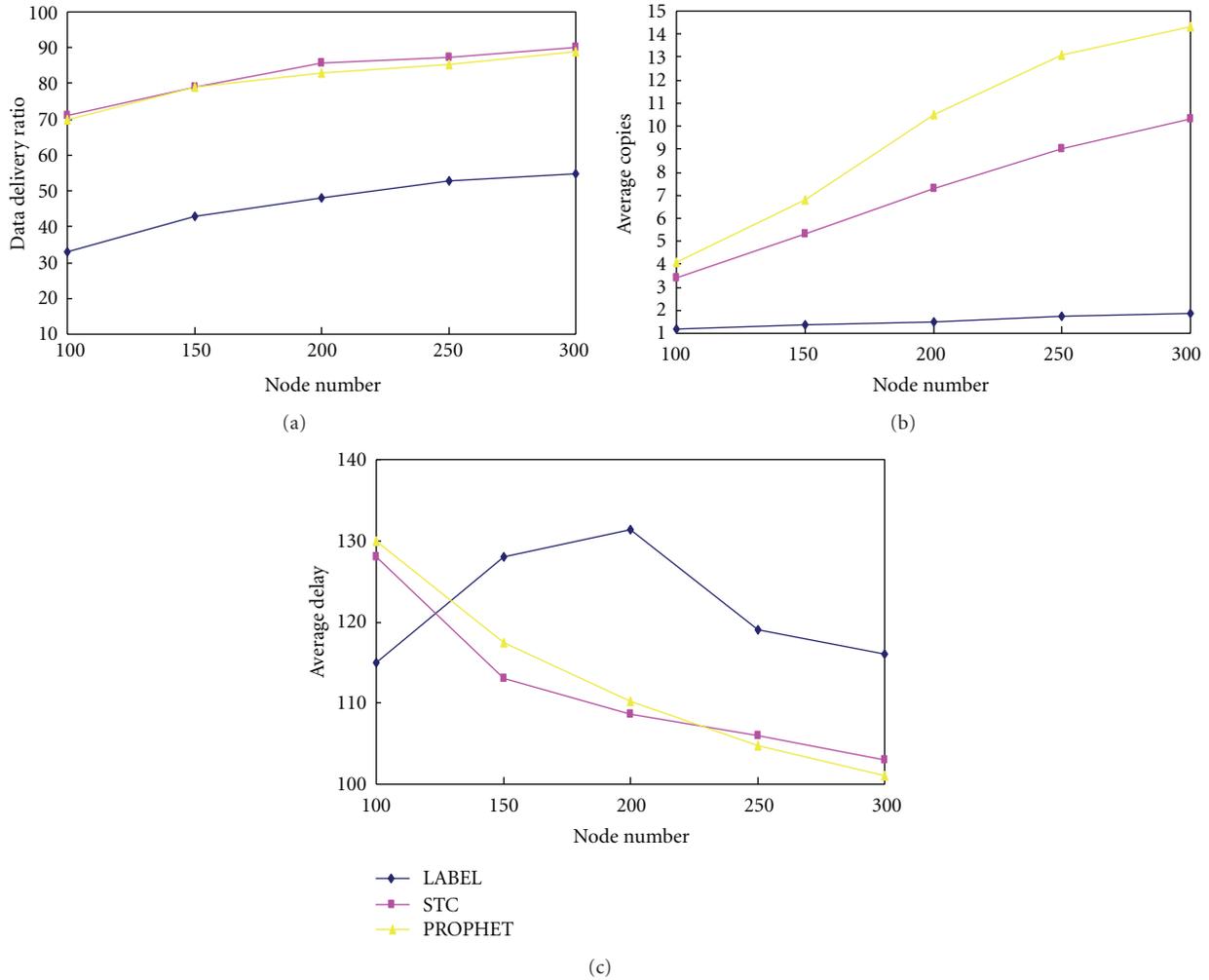


FIGURE 7: Node density's impact on performance. (a) Average delivery ratio. (b) Average copies number. (c) Average delay.

## 7. Conclusions

This paper explores the trip history model (THM) of people's activities and proposes spatial and temporal correlations routing algorithm. In view of the observation that many people have some regularities in their daily movements, we create a model (THM) to describe the mobility characteristics of people who may be treated as mobile nodes and further make motion prediction. Investigation and experiment results have shown that THM can more accurately predict the moving destination of individuals.

In order to make better use of THM, we propose spatial and temporal correlations routing algorithm (STC). Messages route based on the prediction of node movements in THM, and data will always be transferred to the node which has higher transmission probability of meeting the destination node. Combined with judgments of human's trajectory, STC effectively avoid the blindness in data forwarding caused by the random movement model. According to simulation results, STC can obtain a higher data delivery ratio with lower energy consumption and transmission delay.

## Acknowledgment

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

## References

- [1] K. Fall, "A delay-tolerant network architecture for challenged internets," in *Proceedings of the ACM Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications (SIGCOMM'03)*, pp. 27–34, Karlsruhe, Germany, August 2003.
- [2] J. Crowcroft and P. Hui, "Pocket switched networks: real-world mobility and its consequences for opportunistic forwarding," Tech. Rep., Association for Computing Machinery, New York, NY, USA, 2005.
- [3] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Tech. Rep., Duke University, Durham, NC, USA, 2000.
- [4] Y. Wang and H. Y. Wu, "Replication-based efficient data delivery scheme (RED) for delay/fault-tolerant mobile sensor

- network (DFT-MSN),” in *Proceedings of the 4th Annual IEEE International Conference on Pervasive Computing and Communications Workshops*, pp. 485–489, IEEE Computer Society Press, Washington, DC, USA, 2006.
- [5] P. Hui and J. Crowcroft, “How small labels create big improvements,” in *Proceedings of the 5th Annual IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2007*, pp. 65–70, White Plains, NY, USA, March 2007.
- [6] E. M. Daly and M. Haahr, “Social network analysis for routing in disconnected delay-tolerant MANETs,” in *Proceedings of the 8th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc ’07)*, pp. 32–40, Montreal, Canada, September 2007.
- [7] W. J. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy, “Modeling time-variant user mobility in wireless mobile networks,” in *Proceedings of the 26th IEEE International Conference on Computer Communications (IEEE INFOCOM’07)*, pp. 758–766, Anchorage, Alaska, USA, May 2007.
- [8] M. C. González, C. A. Hidalgo, and A. L. Barabási, “Understanding individual human mobility patterns,” *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.
- [9] S. Schonfelder and U. Samaga, “Where do you want to go today?—More observations on daily mobility,” in *Proceedings of the 3rd Swiss Transport Research Conference*, Monte Verita, Ascona, 2003.
- [10] A. Chaintreau, “Impact of human mobility on the design of opportunistic forwarding algorithms,” in *Proceedings of the 26th Annual IEEE Conference on Computer Communications (INFOCOM ’06)*, 2006.
- [11] M. Musolesi and C. Mascolo, “A community based mobility model for ad hoc network research,” in *Proceedings of the 2nd International Workshop on Multi-hop Ad Hoc Networks: From Theory to Reality (REALMAN ’06)*, pp. 31–38, Florence, Italy, May 2006.
- [12] Q. Zheng, X. Hong, and J. Liu, “An Agenda-based Mobility Model,” in *Proceedings of the 39th IEEE Annual Simulation Symposium (ANSS-39-2006)*, p. 8, Huntsville, Ala, USA, 2006.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

