Constructing Time-Varying Contact Graphs for Heterogeneous Delay Tolerant Networks

Xiaoyan Hong, Bo Gu, Yuguang Zeng, Jingyuan Zhang
Department of Computer Science
University of Alabama, Tuscaloosa, AL 35487
{hxy,bgu,zhang}@cs.ua.edu, yzeng3@crimson.ua.edu

***Abstract*—Human mobility, hence the movement pattern of mobile devices, often confines to relatively local geographic areas. Such a movement pattern reduces the opportunities for a message to be disseminated to a more global geographical region using the encounter-based "store-carry-forward" routing approach. On the other hand, different local areas often overlap to cover the entire region. A feasible communication architecture to help message dissemination is to deploy static storage-and-communication devices at those overlapping areas to serve as relays between the local areas. In this paper, we introduce the method to derive the simulation model for this heterogeneous network from contact trace and GPS trace of buses. Our main focus is to model communication properties between the static nodes and the mobile nodes. Typically, they are time-varying link delays formed by a collection of multiple mobile nodes. We further use a Markovian model to describe the time dependency among link delays at each static nodes and use the states to develop a network model for simulation. In the paper, we present simulation results to validate the reproduction of the mobility with the original traces by comparing routing performance. We show that the proposed network model can be used for performance evaluations with inherited realistic. The contributions of this work reside in the reproducibility to the real world traces and its flexibility in configurations. Further, it is the first simulator that enables to produce contact graph for a heterogeneous network with time-varying link properties. Its ability goes beyond simply calculating encounter events, but is well suitable for protocol evaluation in opportunistic networks, mobile social networks and delay tolerant networks.**

I. INTRODUCTION

With mobile devices employing techniques of IEEE 802.11, Bluetooth or other radio solutions, ubiquitous communications can be achieved to help people acquire information at any time and from any place, even without the networking infrastructure. Opportunistic networks enable communications between two mobile nodes when they move into the transmission range of each other. Such contact opportunities among many moving nodes may spread a message among themselves and eventually deliver to the designated destination through "store-carry-forward" routing principle as studied in delay tolerant networks (DTNs). Although many of contact behaviors are sporadic, people's social activities can explicitly or implicitly become driving forces for mobility. The noticeable mobility patterns, if explored, will be of great help for message disseminations in opportunistic networks.

In early work, one important research area is collecting and analyzing movements using wireless contact information and/or GPS traces to help understand the mobile world. The analysis on the contact traces and network accesses records have obtained insightful information of various sorts, such as the distributions of contact time, inter-contact time [15][18], hitting time [16], and network flow characteristics[12]. A recent work [17] developed a method that can translate one trace to many scenarios by capturing the contact information with movement and location constraints. More over, the results provided in Zhang et al.[18] show the periodic behavior of inter-contact times aggregated at routing level for encounter-based data dissemination. Mobility characters on speed and pause time and direction of movement are analyzed in Kim et al.[9]. These contact behaviors have also been explored in many delay tolerant applications [3][10][11].

However, human mobility (hence the movement pattern of mobile devices) often confines to relatively local geographic areas during most of the times. Thus a node will encounter more frequently another node moving within the same local area than one that moves to far areas. In associated with this observation, community based mobility has been studied early [6][13][16]. Such a movement pattern reduces the opportunities for a message to be disseminated to a more global geographical region using the encounter-based approach. On the other hand, we recognize that overlaps of many local areas exist. For example, Figure 1 shows a sample scenario where a few local mobility areas (shaded oval areas) exist and they overlap. A feasible communication architecture to help message dissemination is to deploy static storage-and-communication devices at those overlapping areas. These devices (called message boxes or throw boxes) will be visited by mobile nodes from difference local areas. Thus these boxes serve as relays between two local areas. Optimal deployment of these static devices and efficient routing protocols have been proposed early [4][5][19].

In all, such a mobility phenomena represents a mobile would that is under persistent partition as the result of low density of mobile nodes within each local region and across different regions and also as the result of localized mobility. When considering a graph for this heterogeneous network, namely, a mix of the mobile nodes and the static nodes (at the critical areas), the static nodes serve as the vertices and the collection of mobile nodes form the (virtual) links between the vertices. Due to the dynamic behaviors of mobile nodes that move between the static nodes, the (virtual) links connecting any pair of static nodes show time varying features. In this network model, a mobile node moving between two boxes forms an instance of the link. The traveling time of the mobile node is the link delay. Many mobile nodes moving between the same two boxes form many instances of the link. Thus, the network reveals a stable topology when the message boxes are deployed, yet many time-varying links due to the dynamic occurrences of mobile nodes. For example, two mobile nodes traveling at different time between the same pair of boxes have different delays and lead to different transmission link costs. Figure 1 also shows the communication scenario.

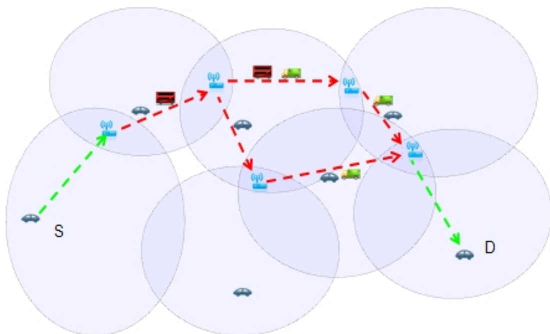


Fig. 1. Local mobility and message dissemination

This mobility phenomena and network graph represent the popular heterogeneous static and mobile network scenarios in the emerging mobile social network applications and novel opportunistic network protocols [2][20]. The evaluation of the new design ideas in these fields calls for a simulation platform that is able to support a realistic hybrid form of mobility model, namely, a movement model for mobile nodes and a placement locations of static nodes; and is also able to capture and support the realistic contact behavior between a mobile node and a static node as well as revealing the time-dependent properties of the (virtual) links. However, the existing trace analysis works do not provide sufficient results to shed light on the dynamic time-varying property of the links, nor the hybrid network composition we described here.

In this paper, we introduce the method to derive the simulation model for the hybrid network from the contact trace of bus-based network [1]. Our analysis starts from abstracting and matching the entities in the contact traces to the communication entities in the hybrid networks; and then deriving communication link properties from the contact traces. Specifically, the static access points and mobile buses from the trace are viewed as the message boxes and mobile nodes respectively. By investigating the link opportunities formed by the buses among the access points (APs), we find that the various link delays can be classified into a limited number of groups, and the link delays also show the cyclic pattern. Therefore we propose a Markovian model to describe the time dependent links for the trace reproduction in the simulation model. In this paper, we also validate the proposed simulation model by evaluating the performance of a single-copy DTN routing protocol in compared with the original trace. The results show that the trace generated simulation model matches nicely to the original trace.

The contributions of this work reside in the reproducibility to the real world traces and its flexibility in configurations. Further, it is the first simulator that is enable to produce contact graph for a hybrid network with time-varying link properties. Its ability goes beyond simply calculating encounter events. The evaluations of the protocols demonstrate that the proposed network simulation model can be well used for protocol evaluation in opportunistic networks, mobile social networks and delay tolerant networks.

The rest of the paper is organized as follows. In Section II, we introduce a communication example and the trace file. Section III introduces the analysis on many link properties. Then, in Section IV, we describe the Markov model that captures the link properties. We also present results that validate the proposed network simulation model. We conclude the paper in Section V.

II. Background

A. Applicable communication example

As mentioned early, Figure 1 shows a sample network scenario. In the figure, the source node sends a message to the destination node by initially leaving the message at the reachable box. Then the message will be carried and disseminated to other boxes through the movements of different mobile nodes. Eventually, the message will be delivered to the box of the home region of the destination and temporally stored there. When the destination node visits the message box, it can collect the stored message.

B. Contact trace

The traces are obtained from UMassDieselNet [1]. The original data collected the contact information of bus-to-bus and bus-to-AP from a bus-based delay tolerant network. The bus-to-bus traces include the IDs of the contact buses, contact times and durations. The bus-to-AP traces are about the contacts between buses and Access Points(AP). There are 40 buses and more than 350 access points in the data set spanning from 10-22-2007 to 11-16-2007. More than three bus routes and four bus shifts including morning, midday, afternoon and evening sub-shifts are found in the trace[18]. We concentrate on the bus-to-AP traces, where the buses moving among the static APs provide the needed mapping from the entities in mobility traces to the communication entities in the hybrid network model.

The format of original contact trace is given in Table I. The two data rows show that the same bus chronologically visits two different APs. Processing the contact data of this moving bus, we obtain the link property between the two APs in Table II. Here, the first visited AP serves as the link source and the other becomes the destination. The time difference between the contact time points indicates the link delay.

TABLE I
TRACE FORMAT

Bus ID	AP MAC address	Contact time	Duration	Location
3030	00:11:5c:03:2e:80	10:7:9	45.438	72.52,42.38
3030	00:0f:90:6a:fa:70	10:10:48	7.675	72.52,42.39

TABLE II
LINK FORMAT

Bus ID	AP as source	AP as destination	Link delay
3030	00:11:5c:03:2e:80	00:0f:90:6a:fa:70	219s

III. TRACE ANALYSIS FRAMEWORK

Using the bus-to-AP traces, we show the one-day route graph for bus 3122 as an example in Figure 2. The MAC address in each oval denotes the visited location (AP). The directed link with a number represents the visiting order and moving direction. As shown in Figure 2, the bus 3122 starts from the AP with address 00:0f:90:51:8c:c0 to the AP with address 00:0f:90:6a:fa:70. Following the visiting order, we find that this bus operates on different routes at consecutive times and connects the corresponding pairs of APs. Further, using multiple one-day traces, we extract the network topology where the vertices are static APs and the (virtual) links are formed by multiple moving buses (shown in Fig. 3). The MAC addresses are used to label vertices. A directed link is generated when there is a record showing a bus moving between two APs. As depicted in Fig.3, most pairs of APs have bidirectional links and the graph becomes strongly connected.

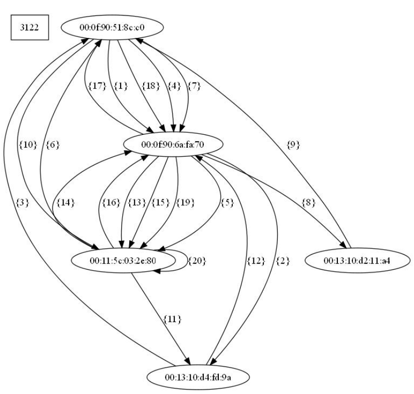


Fig. 2. An example of bus route

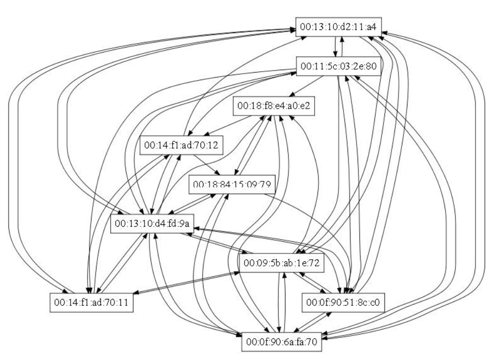


Fig. 3. Contact graph

Our main analysis concentrates on several time-dependent link features. These features include *link inter-arrival time*, *link delay* and *link occurrence sequence*. The *link inter-arrival time* represents the time interval between two consecutive arrival events at the same AP. This is because we view the arrival of a bus at the AP as the emergence of a link. Thus, a shorter inter-arrival time implies a higher contact frequency between the buses and the AP at this specific location. The *link delay* refers to the traveling delay of a bus between two consecutive APs. If a bus could carry messages from one AP to another AP, the delay will be viewed as the message propagation time. At last, the *sequence of bus arrival events* occurring at each AP reflects the sequence of link occurrences, which indicates the potential delay patterns of the stored messages. It can be helpful to predict the link property in the future. In presenting the analysis on these properties, we use examples to illustrate.

A. Link inter-arrival time

From the AP point of view, a contact with a bus is the opportunity for spreading message. A higher contact rate between the AP and the coming buses could have more messages at the AP to be transmitted through the moving buses. So the inter-arrival time becomes a critical factor for the network simulation model. For each AP, we collect the inter-arrival intervals from the bus arrival events and derive the approximated distribution.

Figure 4 shows the probability density of the inter-arrival time and the fitted distributions for AP 3 as an example. The blue box represents the probability density of the discrete time data and the red curve gives the fitted probability density function. The histogram shows that a single distribution model can not capture the distribution patterns. In fact, the previous works [18] has pointed out that the distribution tends to be a mixed normal distribution. Here, we also use several normal distributions to approximate the time data. The parameters of the distributions are given in the figures. In addition, from the realistic point of view, it is reasonable to accept the factor that the inter-contact time by a bus at one AP follows normal distribution due to the dynamic road traffic conditions that cause the variance. And the mixed normal distribution comes from the many buses visiting the same AP.

Figure 5 provides the sequences of buses arriving at AP 3 at different time points. The total time span of the trace is divided into equal-sized intervals based on the average inter-arrival time. The intervals are numbered as steps. The buses are color coded and the number of buses arrived during each interval is shown. Many buses visited AP 3 with different patterns. We can explain that the resultant distribution combines independent mixed distributions.

B. Link delay

The link delay distribution at an AP is derived from the traveling delay of a bus starting from that AP. The trace format for extracting the link delay is shown in Table II. Figure 6 gives the probability distribution of link delay for AP 3. Again, the mixed normal model is used to fit the data of link delay. From the figure, we observed that the irregular data could be approximately clustered into several groups with different typical parameter values. The fitting accuracy can be enhanced by increasing the number of clusters. The typical parameter value of each cluster is set to be the mean of normal distribution.

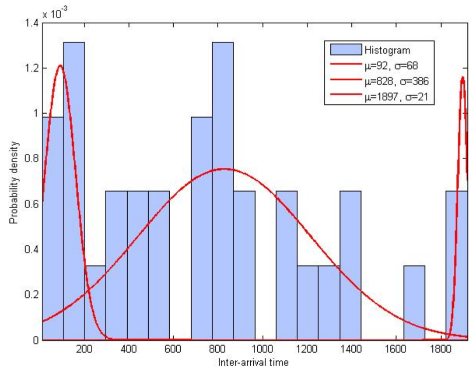


Fig. 4. Distribution of inter-arrival time at AP 3(00:0f:90:51:8c:c0)

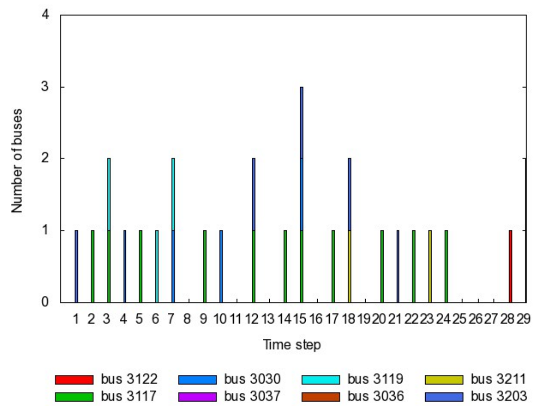


Fig. 5. Bus occurrences at AP 3 (00:0f:90:51:8c:c0)

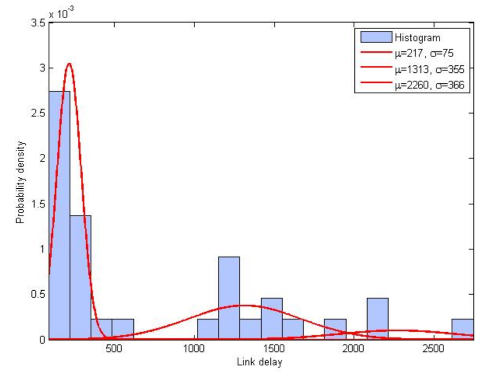


Fig. 6. Distribution of link delay at AP 3(00:0f:90:51:8c:c0)

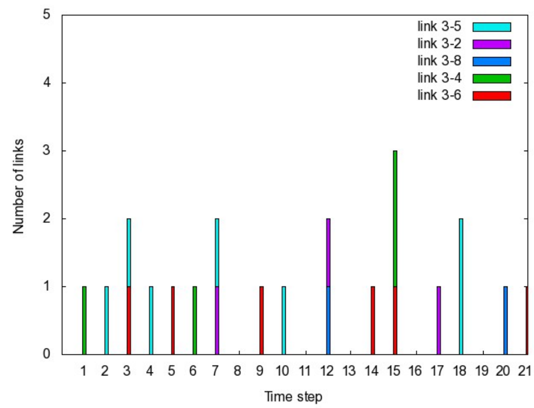


Fig. 7. Link occurrences at AP 3

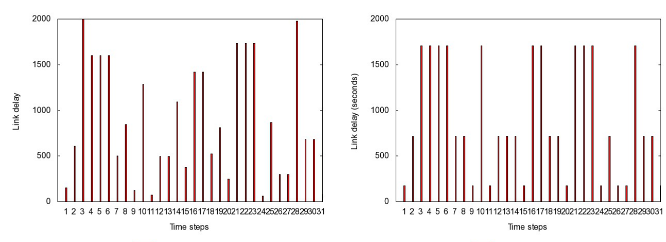
C. Link occurrence sequence

The knowledge about the link occurrence sequence at a AP helps to study the changing pattern of the link delay at the AP, which helps in predicting the future link properties. The interval for collecting the link sequence is derived from the mixed distribution of the link inter-arrival time. During the interval, the involved link (the source-destination pair), the sequence number of the occurrence and the delay are recorded. Fig.7 shows the sequences occurring at AP 3 during each expected inter-arrival period. In addition, the figure also shows the occurrence patterns. In Fig.7, there are total five outgoing links from AP 3. And the links have different frequencies of occurrences.

Further, we show the link delays corresponding to the link sequence at the AP 4 in Fig. 8(a). While original delay values are shown in Fig. 8(a), we show that these link delays can be approximately grouped into several categories. The results after numerical classification are given in Fig.8(b). These processed results imply that the occurrence sequence can be classified based on the changing pattern of link delay. Such feature helps us to use a stochastic model to describe the time dependency in link properties.

D. Summary

The three link properties are essential in modeling the time-varying links in the hybrid delay tolerant network. Firstly, the values of the link delay can serve as the main metric to categorize the link occurrence pattern. As shown in the analytical results, the values of link delay reveal the clustering feature. It is possible to extract the typical value to represent a batch of link arrivals. In addition, the expected link inter-arrival time could represent the transition interval between a group of links with their typical clustered delays. After that, the sequence of link occurrences can further characterize the transitional pattern between the different links, which could be used to predict the link delay in next time interval. The results obtained through the trace analysis contribute to develop the



(a) Link delay sequence at AP 4

(b) Grouped link delay sequence at AP 4

Fig. 8. Link property and Markov model

simulation model for the targeted network communications.

IV. Network Simulation Model

A. Markov chain based network model

In order to model the time-dependent links between the message boxes that are formed collectively by the mobile nodes, a Markov chain model is proposed to represent the dynamic delay states at a static message box. Each state is associated with a link delay. At each massage box, there will be a set of link states describing the evolution of the link delays. Denote S^k be the set of link states for message box k. We use K-mean clustering algorithm to classify all the delay samples into a limited number of groups [14]. After trying multiple K values, we obtain the best set of recommended number of groups and their corresponding delays for the states. The state transition interval equals to the expected link inter-arrival time. In order to derive the state transition, we use the data of the sequences of the link delay to get the frequency. Let s_{ij} denote the frequency from state i to state j, where $i, j \in S^k$. The probability of state transition from state i to state j is $p_{ij} = \frac{s_{ij}}{\sum_{k \in S} s_{ik}}$.

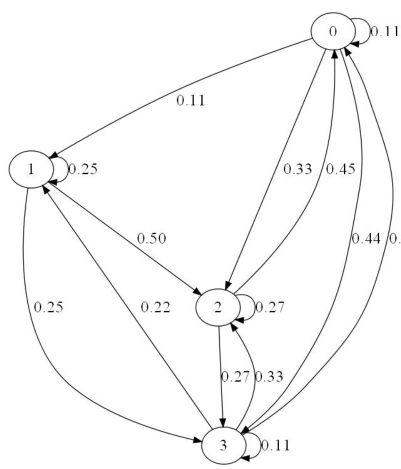


Fig. 9. State transition graph at box AP 3

Fig.9 gives the state transition graph of the message box AP 3. The data values are derived from the trace. In the graph, there are total four states with the corresponding delay values. Table III provides more details about the link information including the link destination and link delay for each state. If there is no bus arriving at the AP during the interval, the link delay of the special state equals to 10000s in the model. State 2 in Table III denotes such special state and the link destination is set to be NULL.

B. Network model validation

The proposed Markov chain based network simulation model (MC) can characterize the time-dependent virtual link properties and be used to generate the contact trace for extensive simulation. We develop our new simulator out from the DTN Simulator ONE [8]. In order to validate our proposed simulation model, we compare the trace generated from it with the original trace in terms of the routing performance. Since we focus on the validation of the link events between a pair of message boxes, the single-copy message routing strategy is suitable for validation. The routing protocol chosen is First Contact routing (FC) mentioned in [7]. In FC routing, only one message copy exists in the network and message carrier hands over the message to the first encountered node. In our case, encounter happens between a static node and a mobile node. The main metric for comparison is the delivery delay. We vary the network traffic generation rates, and use these two traces as the mobility model. The results are presented as the cumulative distribution function of the delivery delay of each message.

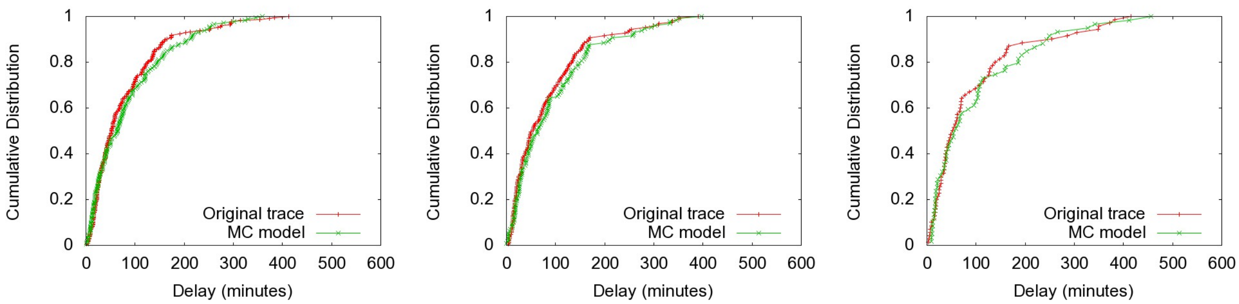
TABLE III
Link delays for different states at box AP 3

State	Destination	Link delay
0	2	996
0	5	983
0	6	1414
0	8	547
1	5	1181
1	6	2074
2	NULL	10000
3	2	205
3	4	237
3	5	220
3	6	136
3	8	339

Fig. 10(a) gives the comparison CDF results between the original trace and the proposed Markov based trace. The message generation rate is 1/60 and each message is 50KB. As shown in Fig. 10(a), the distribution of MC model trace is close to that of original trace although small gaps exist between two curves at a few places. The first reason for the difference between the two curves is that we use the expected value to approximate all the delays in each state of the Markov model. The second reason is that the proposed model treats the outgoing links (to the neighboring boxes) during each state interval uniformly, but the link occurrence sequence has shown that the distribution could be arbitrary in the original trace. Fig. 10(b) and Fig. 10(c) further provide the cumulative distributions of the delivery delay with the smaller message generating rates. The two figures show similar matchings between the original trace and the generated trace.

V. Conclusion

In this paper, we motivated the need for creating a simulation model for a novel heterogeneous static and mobile networks, where the time-varying delay properties are the dominating network features. We investigated the delay properties such as inter-arrival time, link delay and link occurrence sequence from the bus traces obtained from a bus-based network. We propose a Markov model to describe the link delays and their changing patterns. We also used the analysis results to generate the synthetic network model and validated it by comparing with the original trace through routing performance. The results show that the proposed network model can be used for performance evaluations with inherited realistic.



(a) Message generating rate 1/60

(b) Message generating rate 1/80

(c) Message generating rate 1/200

Fig. 10. Delivery delay at different message generating rates

It can be well suitable for protocol evaluation in opportunistic networks, mobile social networks and delay tolerant networks.

REFERENCES

[1] A. Balasubramanian, Y. Zhou, B. N. Levine, A. Venkataramani, and B. Croft. Crawdad trace umass/diesel/transfer/ap_connectivity v.2007-12-02. Downloaded from http://crawdad.cs.dartmouth.edu/umass/diesel/transfer/ap_connectivity, Dec. 2007.

[2] M. Chuah, Y. Peng, and P. Hui. Cooperative User Centric Information Dissemination in Human Contact-Based Networks. In *IEEE 2nd International Workshop on P2P Networking*, Dec, 2010.

[3] W. Gao, Q. Li, B. Zhao, and G. Cao. Multicasting in delay tolerant networks: a social network perspective. In *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, 2009.

[4] B. Gu and X. Hong. Optimal Routing Strategy in Throw-box based Delay Tolerant Network. In *CHINACOM 2011*, Harbin, China, August 2011.

[5] B. Gu and X. Hong. Capacity-Aware Routing Using Throw-Boxes. In *IEEE Global Telecommunications Conference*, Houston, Texas, USA, Dec., 2011.

[6] P. Hui, E. Yoneki, S. yan Chan, and J. Crowcroft. Distributed Community Detection in Delay Tolerant Networks. In *ACM Sigcomm Workshop MobiArch*, 2007.

[7] S. Jain, K. Fall, and R. Patra. Routing in a Delay Tolerant Network. In *Proc. ACM Sigcomm*, pages 145–158, 2004.

[8] A. Keränen, J. Ott, and T. Kärkkäinen. The ONE Simulator for DTN Protocol Evaluation. In *Proceedings of the 2nd International Conference on Simulation Tools and Techniques*, New York, NY, USA, 2009. ICST.

[9] M. Kim, D. Kotz, and S. Kim. Extracting a mobility model from real user traces. In *INFOCOM 2006. 25th IEEE International Conference on Computer Communications. Proceedings*, pages 1 –13, 2006.

[10] A. Lindgren, A. Doria, and O. Schelen. Probabilistic routing in intermittently connected networks. *SIGMOBILE Mobile Computing and Communications Review*, 7(3), 2003.

[11] C. Liu and J. Wu. Routing in a cyclic mobispace. In *Proceedings of the tenth ACM international symposium on Mobile ad hoc networking and computing*, 2008.

[12] Y. Lou and Y. Yin. A decomposition scheme for estimating dynamic origindestination flows on actuation-controlled signalized arterials. *Transportation Research*, 18(9):643–645, 2010.

[13] M. Musolesi and C. Mascolo. A community based mobility model for ad hoc network research. In *REALMAN '06: Proceedings of the 2nd international workshop on Multi-hop ad hoc networks: from theory to reality*, 2006.

[14] D. T. Pham, S. S. Dimov, and C. D. Nguyen. Selection of K in K-means clustering. In *Proc. IMechE*, volume 219, pages 103–119, 2005.

[15] P. Samar and S. B. Wicker. Link Dynamics and Protocol Design in a Multihop Mobile Environment. *IEEE Transactions on Mobile Computing*, 5(9):1156–1172, 2006.

[16] K. P. Wei-jen Hsu, T. Spyropoulos and A. Helmy. Modeling Time-Variant User Mobility in Wireless Mobile Networks. In *IEEE INFOCOM*, 2007.

[17] J. Whitbeck, M. D. de Amorim, and V. Conan. Plausible mobility: inferring movement from contacts. In *Proceedings of the Second International Workshop on Mobile Opportunistic Networking MobiOpp '10*, 2010.

[18] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, MobiCom '07, pages 195–206, New York, NY, USA, 2007. ACM.

[19] W. Zhao, Y. Chen, M. Ammar, M. D. Corner, B. N. Levine, and E. Zegura. Capacity Enhancement using Throwboxes in DTNs. In *IEEE MASS*, 2006.

[20] G. Zyba, G. M. Voelker, S. Ioannidis, and C. Diot. Dissemination in Opportunistic Mobile Ad-hoc Networks: the Power of the Crowd. In *IEEE Infocom, Shanghai, China, April 2011, pages 1179-1187.*, Shanghai, China, 2011.