

# Location-aware Routing for Delay Tolerant Networks

Ye Tian

Anhui Province Key Laboratory on High Performance Computing  
School of Computer Science and Technology  
University of Science and Technology of China  
Hefei, Anhui 230027, China  
Email: yetian@ustc.edu.cn

Jiang Li

Department of Systems and Computer Science  
Howard University  
Washington DC 20059, USA  
Email: lij@scs.howard.edu

**Abstract**—In this paper, we sought to understand the reasons causing failures and delays of message delivery in Delay Tolerant Networks (DTN), and to use this understanding for improving the network. By studying two real-world datasets, we found that node isolation is prevalent, which largely accounts for the inefficiencies in DTN’s message delivery. In addition, by analyzing nodes’ contact-location relationship, we found that individual and system-wide key locations exist and their existence suggests potential improvements. Motivated by our observations, we designed a location aware routing scheme for DTN networks. With simulation-based experiments, we compared our proposal with other representative DTN routing schemes, and showed that with the awareness of the location information, our solution can deliver more messages within shorter delays, therefore improves the network intensively.

## I. INTRODUCTION

With advance of wireless technologies and prevalent usage of portable wireless devices, in recent years, the idea of Delay Tolerant Networks (DTN[1]) has been proposed. Simply speaking, a DTN network is an ad-hoc network formed by portable devices such as cell phones or PDAs, and in such a network, device-to-device communication is enabled using the *store-and-forward* paradigm. As the only communication opportunity in a DTN network is the device contact caused by uncontrollable human mobility, it is essential to understand and exploit the human contacts in designing a DTN routing scheme.

Many routing schemes have been proposed for DTN networks in recent years. However, for most of them (e.g. [2][3][4][5][6]), a “flat” network was assumed, where each node plays an equal role. In these schemes, decisions on message forwarding are based on some destination dependent quality metrics, and the message is forwarded towards the nodes with better qualities. On the other hand, a recent study[7] shows that social structure of the network exist and should be labeled and exploited. In our work, we also labeled the inherited structure of the network explicitly, but unlike the previous work, we labeled the nodes based on their location

This work is funded by the Specialized Research Fund for the Doctoral Program of Higher Education of China 20093402120020, and was supported in part by US NSF grant CNS-0832000 and the Mordecai Wyatt Johnson Program at Howard University.

visiting information as well as their contact information to make forwarding decisions.

Two contributions were made in this paper: first of all, we conducted an insightful analysis on the contact and location information of nodes in DTN networks. By examining two real-world datasets, we found that the phenomenon of node isolation is prevalent and is the major reason for the inefficiencies in DTN’s message delivery. We also found that there exist special locations for individual node and some of them are of system-wide importance. Motivated by these observations, we proposed a location aware routing scheme for DTN networks. In addition, we demonstrated the superiority of the proposed scheme over other representative schemes via extensive simulation experiments using the contact trace from real-world dataset.

The remainder part of this paper is organized as follows: related works are surveyed in Section II; we introduce and analyze the real-world datasets in Section III; in Section IV, we propose the location aware routing scheme for DTN networks; performance evaluation and comparison are presented in Section V; finally, Section VI concludes this paper.

## II. RELATED WORK

Early works on the DTN routing problem were focused on deploying and using dedicated nodes, such as static throwboxes[8] and controllable mobile ferries[9]. In recent years, DTN networks composed of portable devices carried by human beings have received great attentions. For message routing in such a network, although the epidemic scheme[10] based on flooding makes use of every device contact, it is regarded as impractical due to the large volume of message duplicates generated. The Spray and Wait routing[11] made an effort on reducing the duplicates by assigning limitations on message copies. Besides the epidemic-style ones, for majority of the DTN routing schemes, messages are forwarded towards the nodes which are considered as better candidates for delivery. Examples include MED[2], FRESH[3], PRoPHET[4], MobySpace[5], SimBet[6], and delegation forwarding[12]. These schemes vary on their criteria in evaluating the next hop, and for most of them, the criteria is destination dependent. Besides simply comparing destination dependent metrics and

forwarding messages on a “flat” network, inherited social structure of the network is also exploited. In the recent proposed BUBBLE Rap routing scheme[7], communities among the nodes are detected and explicitly labeled, and different forwarding strategies are applied in different routing phases based on the community structure.

### III. ANALYSIS OF REAL-WORLD DATASETS

#### A. Datasets under study

To investigate the device contacts in DTN networks, we selected two real-world datasets which contain long-time mobility and contact information from a large number of participating nodes for our study. The two datasets are from the MIT Reality Mining project[13] and from the UCSD Wireless Topology Discover project[14] in respective. For details of the datasets, please refer to our technical report[15]. In particular, in the UCSD dataset, we considered a contact between two devices happened if they were associated with a same AP simultaneously, as in the previous works [16][17]. In the remainder part of this paper, we refer to the two datasets as *Reality* and *UCSD* for simplicity.

#### B. Connectivity and isolation

TABLE I  
COMPONENTS AND ISOLATION IN CONTACT GRAPH

Contact graph	Num. of comp.	Num. of isolated nodes	Largest comp.	2 <sup>nd</sup> largest comp.
$G(Reality, 25)$	13	12	84	1
$G(Reality, 35)$	15	14	82	1
$G(Reality, 45)$	23	21	73	2
$G(UCSD, 10)$	89	86	155	2
$G(UCSD, 15)$	108	106	137	2
$G(UCSD, 20)$	128	125	58	54

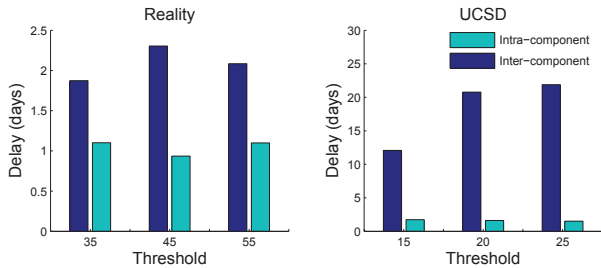


Fig. 1. Comparison of intra- and inter-component delays for Reality and UCSD

For DTN networks, how to avoid the message delivery failure and reduce the delay is the major concern. In this section, we sought to find the fundamental reasons causing these inefficiencies by studying *contact graph* of the datasets above introduced.

Given a dataset  $d$ , a contact graph  $G(d, t) = (V, E)$  is defined as an undirected graph, with  $V$  as the vertex set where each node in the dataset corresponds to a vertex. For any two vertexes  $v_i$  and  $v_j$ , there is an edge  $(v_i, v_j)$  in the edge set  $E$  if the number of the contacts between the two corresponding nodes in  $d$  is no smaller than a threshold number  $t$ . By

varying contact thresholds, we obtained a number of contact graphs from Reality and UCSD, that is,  $G(Reality, t)$  with  $t = 25, 35, 45$  and  $G(UCSD, t)$  with  $t = 10, 15, 20$  in respective. The summaries of the contact graphs are listed in Table I. From the table one can see that nodes are heterogeneous regarding their connectivity. In Reality, one *giant component* containing more than half of the nodes exists, and nearly all the other nodes are either *isolated* or form *tiny components* (a component contains very few nodes). When the threshold gets increased, more nodes are isolated from the giant component but the giant component still exists. For UCSD, similar observations could be made except that with the increase of the threshold, one giant component splits into two giant components of approximately the same size. In the following part of this paper, we refer to a dataset’s giant component as *component*, and refer to the isolated nodes and the nodes in tiny components together as *isolated* nodes. Clearly for any node in the dataset, it is either in a component or be isolated.

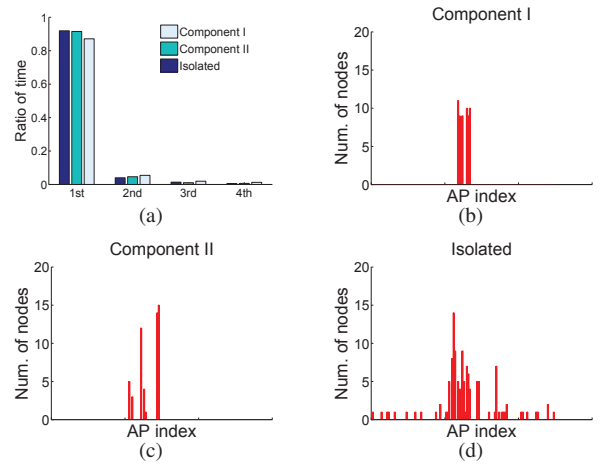


Fig. 2. (a) Comparison of APs in ratio of time a node stayed; (b-d) Histograms of times an AP became home location of nodes in Component I, of nodes in component II, and of isolated nodes

With node isolation widely observed, to investigate its implication on the DTN routing problem, we studied two metrics of a contact graph, namely the *intra-component delay* and the *inter-component delay*. To define the metrics, we modified the contact graph into a weighted contact graph  $G^w(d)$ , where for each edge on the graph, its weight is defined as the averaged inter-contact time between the two corresponding nodes. With the weighted contact graph, the intra-component delay is defined as the mean distance between any two nodes in a same component using the MED algorithm[2], and the inter-component delay is defined as the mean distance between one component and one isolated node.

We show the calculated intra- and the inter-component delays under different thresholds in Fig. 1. From the figure, one can see that for both datasets, the inter-component delays are much longer than the intra-component ones. In fact in practical DTN routing, failures are more likely to be expected than the long delays due to timeout. The observation suggests

that node isolation is the major reason for the latency and the failure in DTN’s message delivery, and should be our focus in designing the routing scheme.

### C. Contact-location relationship

In this section, we studied the locations where nodes make their contacts. As location information is not available in Reality, we only used the UCSD dataset. We set a threshold of 20 contacts for detecting the components, and as listed in Table I, we had 58 nodes in one component (referred to as Component I), and 54 nodes in the other component (referred to as Component II), and all the other 133 nodes were regarded as isolated.

1) *Home location*: To find a specific node, the most direct way is to identify the location where it is most likely to stay. Therefore we considered the location (i.e., AP) each node stayed for the longest time accumulatively in UCSD. Fig. 2(a) shows the ratios of the time nodes in different groups spent at their most favorite locations. We also identified the 2nd, 3rd, and 4th favorite locations of the nodes and plotted the ratios on the figure. It is surprising to see that a node spends more than 90% of its time at just one location, regardless whether it is in a component or isolated. We refer to this location as the node’s *home location*.

With each node’s home location identified, we also counted how many times an AP becomes a node’s home location and plotted the histograms over all the APs for each group in Fig. 2(b-d) respectively. One can see from the histograms that although some nodes share their home locations, nodes are in general not having a common home location, especially for the isolated ones. Our observation suggests that although nodes tend to stay at their home locations, none of them is of a system-wide importance.

and 4th popular APs for comparison. The ratios of the nodes a node in different group could meet at these locations are calculated and plotted in Fig. 3(a). From the figure one can see that a node could encounter over 70% of the total nodes it could ever meet at just one location, regardless whether it is in a component or isolated. We refer to this location as the node’s *individual pub location*.

We also counted the times an AP is chosen as the individual pub location by nodes in different groups and plot the histograms in Fig. 3(b-d). It is surprising to see that overall there are nearly half of the nodes having one location, i.e. AP 173, as their pub location, and AP 173 is the most and the only popular pub location candidate for both component nodes and isolated nodes. For such a location which is the individual pub location of a large population, such as AP 173 in UCSD, we name it as a *system-wide pub location*<sup>1</sup> as it is system-wide important. Meanwhile, we also found that in general a node is not likely to have a same location as its home and pub locations.

As many nodes were observed to go to the system-wide pub locations, we investigated their visiting patterns to this location. By analyzing the nodes in UCSD with AP 173 as their pub locations, we found that 1) there is only a limited number of nodes which visited AP 173 frequently; 2) a node did not meet many other nodes in its one single visit to AP 173; 3) most pairs of nodes did not meet at AP 173 repeatedly. We omit the detailed statistical results here for space reason, interested readers can refer to our technical report [15]. Finally, our observation show that although pub locations of system-wide importance exist, they are not hot locations for each single node’s activity in DTN networks, therefore to exploit them, a well designed routing scheme is required.

## IV. ROUTING SCHEME

Motivated by the observations made in the previous section, we propose a message routing scheme for DTN networks to avoid the failures and reduce the delays in message delivery. In our scheme, we assume that the network area is divided into cells called locations, components among the nodes are identified and nodes’ pub location information as well as their belongings to the components are available. For the concept of component in our routing scheme, other definitions (such as the “community” defined in [18]) can be used as long as frequently contacting nodes are grouped together.

We proposed a location aware routing scheme (referred to as *Location*) under the DTN networking environment. In the Location routing, we differentiate the operations of *forwarding* a message and *replicating* a message. As in many message forwarding algorithms (e.g., [7], [19]), a node holding a message can forward the message to a non-destination node only once, after that, the node becomes forwarding inactive and can only pass the message to the destination. On the other hand, a node holding the message can replicate it to other nodes without such constraint.

<sup>1</sup>We will use the term “pub location” for the individual and system-wide pub location depending on the context.

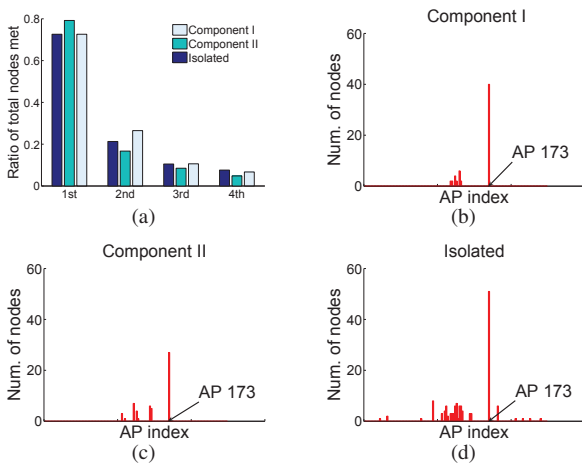


Fig. 3. (a) Comparison of APs in ratio of nodes a node could meet; (b-d) Histograms of times an AP became pub location of nodes in Component I, of nodes in Component II, and of isolated nodes

2) *Pub location*: Meanwhile, we also considered the location where a node could encounter many other nodes. For each node in UCSD, we identified the AP where it could meet the largest number of distinct nodes, as well as its 2nd, 3rd,

---

**Location Assisted Routing**

---

**PhaseIRoute** (*Current*)

```

1 foreach Encountered do
2   if (Encountered.Comp = Dest.Comp)
3     Forward(Encountered)
4     PhaseIIRoute(Encountered)
5   if (Current = Source && Is_New(Encountered.PubLoc) =
true)
6     Replicate(Encountered)
7     PhaseIRoute(Encountered)
8   if (Encountered.PubLoc = Current.PubLoc &&
Encountered.NodesMetAtPub > Current.NodesMetAtPub)
9     Forward(Encountered)
10    PhaseIRoute(Encountered)

```

**PhaseIIRoute** (*Current*)

```

1 foreach Encountered do
2   if Encountered.Metrics > Current.Metrics
3     Forward(Encountered)
4     PhaseIIRoute(Encountered)

```

---

Fig. 4. Location aware routing scheme

The message routing procedure in Location is divided into two phases: the first phase happens when the forwarding active node is outside the component of the destination or the destination node is isolated; the second phase happens when the forwarding active node is in the same component of the destination. A brief description of the scheme can be found in Fig. 4.

In the second phase, many message routing strategies such as the ones used in MED[2], FRESH[3], PROPHET[4], and MobySpace[5] can be applied. The basic idea is to compare certain destination dependent quality metrics and forward the message to a node with better quality. For our experimental evaluation, we chose the *Greedy* strategy used in [19]. We stress that our focus here is not on comparing the metrics based routing strategies in the second routing phase, but is on the first routing phase (i.e., routing outside the destination’s component).

In the Location routing scheme, we exploited the location information as well as the social network structure (i.e. components) in the first routing phase. In detail, if a forwarding active node outside the destination’s component encounters a node, it examines whether the encountered node is in the same component of the destination, and adopts different strategies as described in line 2-4 and line 8-10 respectively. In addition, for the source node, if it encounters a node with a pub location it has not seen before, it replicates the message on it, and the replicated node starts to forward the message among the nodes with the same pub location as described in line 5-7.

## V. PERFORMANCE EVALUATION

In this section, we evaluated the performance of the proposed location-aware approach, and compared it with other four representative DTN routing schemes, which are named *Epidemic*, *Wait*, *Greedy* and *Social*. In particular, the Social routing scheme applies exactly the same strategy used in BUBBLE Rap[7] when routing messages outside of the destination’s component. For details of these schemes, please refer to our technical report[15].

For each routing scheme under study, given a message delivery task between a source and a destination node within a delay constraint, we are interested in 1) whether the message can be successfully delivered within the constraint; 2) how many times the message is copied and transferred during the routing procedure. The former metric regards the effectiveness of the routing scheme while the latter indicates the cost of bandwidths and storages paid for delivering the message.

### A. Overall comparison

TABLE II  
COMPARISON OF ROUTING SCHEMES USING UCSD TRACE

	Message delivered	Message copies
Epidemic	76.0±8.2	104.6±5.92
Wait	25.2±2.0	1.0±0.0
Greedy	52.6±3.3	3.0±0.25
Social	63.0±5.5	6.7±0.4
Location	67.0±5.1	12.1±1.2

We evaluated and compared performances of the five schemes above mentioned using the contact trace from the UCSD dataset. An event-driven simulator is developed. In our first experiment, for each routing scheme, we randomly selected one hundred source/destination pairs for message delivery, and we did not restrict the delay constraint. The averaged number of successfully delivered messages and the averaged number of message copies are listed in Table II.

From the table one can see that Epidemic can deliver the largest number of messages, while Wait delivers the fewest. For the three forwarding-style schemes, i.e., Greedy, Social, and Location, Location delivers more messages, suggesting its superiority over the other two schemes. For the message copies, Epidemic duplicates and transfers much more copies than the other schemes, making it impractical under the context of DTN networks. Among the three forwarding-style schemes, Location has more copies than the other two. However, the message copies generated by Location are much fewer than by Epidemic, and under moderate conditions, Location could be considered as real-world practical. As Social and Location perform much better than Greedy, in our following study, we only focus on comparing the two schemes. In addition, as both schemes just incur moderate costs, in the following study we only focus on their effectiveness (i.e. messages successfully delivered).

### B. Detailed study

In Section III, we point out that node isolation is the main reason for message delivery failures and delays, we testify this argument in this experiment. Instead of randomly selection, four cases of source/destination pairs are considered: in the “Component → Component” case, the source and destination nodes are in different components; in the “Component → Isolated” case, we select the source node from a component, and select an isolated node as the destination; the third and the fourth cases are referred to as “Isolated → Component” and “Isolated → Isolated”, with denotations of the same meanings.

In addition, for each case we only select the infrequent visitors to its pub location.

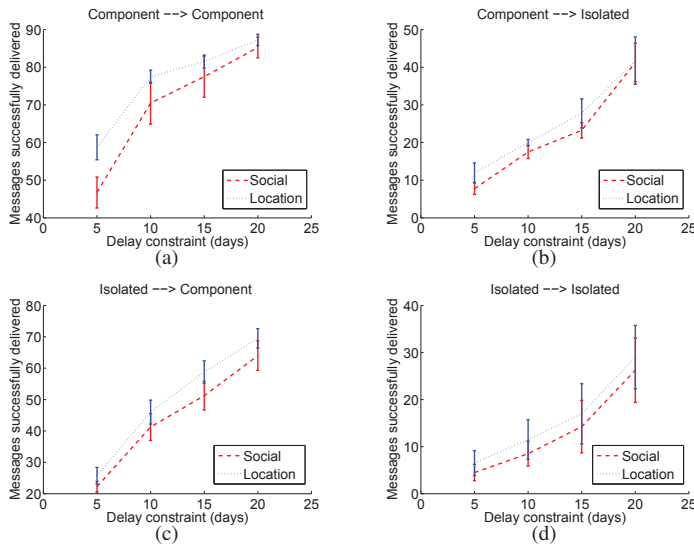


Fig. 5. Number of successfully delivered messages by Social and Location with varying delay constraint under different source/destination cases using UCSD trace

We randomly chose one hundred source/destination pairs for each case for simulation. In this experiment, a delay constraint was imposed as messages must be delivered within this delay. For Social and Location routing schemes under each source/destination case, we varied the delay constraint and plotted the successfully delivered messages in Fig. 5. From the figure, first of all one can see that performances of the two routing schemes under different source/destination cases differ greatly: for both schemes, they have the best performance under the “Component  $\rightarrow$  Component” case, but in cases where isolated nodes are involved, either as source or as destination, fewer messages could be delivered, especially when the destination is isolated. The observation here confirms our argument that node isolation is the major reason for failures and delays in DTN’s message delivery. Moreover, by comparing the two routing schemes, one can see that although the Location scheme suffers node isolation as well, it outperforms Social all the time. In particular, under “Component  $\rightarrow$  Component”, “Component  $\rightarrow$  Isolated”, and “Isolated  $\rightarrow$  Isolated”, when the delay constraint is stringent, say, 5 days, Location delivers much more messages than Social. Apparently, this feature makes Location an attractive solution under environments where message deliveries are not so “delay tolerant”. Finally, for all the four cases, it is observed that Location consistently delivers more messages than Social. We believe this is because by the replicating operation in the Location scheme, the source node has more chances to forward the message into the destination node’s component.

## VI. CONCLUSION

In this paper, we considered the message routing problem under Delay Tolerant Networks (DTN). By analyzing real-world datasets, we found that node isolation is prevalent,

and it is the major reason for message delivery failures and long delays in DTN networks. We also investigated the nodes’ contact-location relationship, and found that there exist pub locations which are of system-wide importance.

Motivated by these observations, we considered incorporating the location information of nodes as well as their contact information in routing messages by designing a location aware routing scheme for DTN networks. We demonstrated the scheme’s effectiveness by comparing it with a number of representative solutions via simulation-based experiments, where real-world contact traces is used to drive the simulator. In particular, we show that with the awareness of contact-location relationship, message delivery jobs involved with isolated nodes could be better accomplished, comparing with the schemes without such concern (e.g., Bubble Rap[7]).

## REFERENCES

- [1] Delay Tolerant Networking Research Group. [Online]. Available: <http://www.dtnrg.org/>
- [2] S. Jain, K. Fall, and R. Patra, “Routing in a delay tolerant network,” in *Proc. of ACM SIGCOMM’04*, Portland, OR, USA, Aug. 2004.
- [3] H. Dubois-Ferriere, M. Grossglauser, and M. Vetterli, “Age matters: efficient route discovery in mobile ad hoc networks using encounter ages,” in *Proc. of ACM MobiHoc’03*, Annapolis, MD, USA, Jun. 2003.
- [4] A. Lindgren, A. Doria, and O. Schelen, “Probabilistic routing in intermittently connected networks,” *ACM SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 7, no. 3, pp. 19–20, 2003.
- [5] J. Leguay, T. Friedman, and V. Conan, “Evaluating mobility pattern space routing for dtn,” in *Proc. of IEEE INFOCOM’06*, Barcelona, Spain, Apr. 2006.
- [6] E. Daly and M. Haahr, “Social network analysis for routing in disconnected delay-tolerant manets,” in *Proc. of ACM MobiHoc’07*, Montreal, Canada, Sep. 2007.
- [7] P. Hui, J. Crowcroft, and E. Yoneki, “Bubble Rap: social-based forwarding in delay tolerant networks,” in *Proc. of ACM MobiHoc’08*, Hong Kong, May 2008.
- [8] W. Zhao, Y. Chen, M. Ammar, M. Corner, B. Levine, and E. Zegura, “Capacity enhancement using throwboxes in dtms,” in *Proc. IEEE Intl Conf on Mobile Ad hoc and Sensor Systems (MASS’06)*, Vancouver, Canada, Oct. 2006.
- [9] W. Zhao, M. Ammar, and E. Zegura, “Controlling the mobility of multiple data transport ferries in a delay-tolerant network,” in *Proc. of IEEE INFOCOM’05*, Miami, FL, USA, Mar. 2005.
- [10] A. Vahdat and D. Becker, “Epidemic routing for partially-connected ad hoc networks,” Duke University, Tech. Rep. CS-200006, 2000.
- [11] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, “Spray and wait: an efficient routing scheme for intermittently connected mobile networks,” in *Proc. of ACM SIGCOMM Workshop on Delay-Tolerant Networking (WDTN’05)*, Philadelphia, PA, USA, Aug. 2005.
- [12] V. Erramilli, M. Crovella, A. Chaintreau, and C. Diot, “Delegation forwarding,” in *Proc. of ACM MobiHoc’08*, Hong Kong, May 2008.
- [13] MIT Reality Mining Project. [Online]. Available: <http://reality.media.mit.edu/>
- [14] UCSD WTD Project. [Online]. Available: <http://sysnet.ucsd.edu/wtd/>
- [15] Y. Tian and J. Li, “Location-aware routing for delay tolerant networks,” Univ. of Science and Technology of China, Tech. Rep., 2010. [Online]. Available: <http://staff.ustc.edu.cn/~yetian/pub/DTNRouting.pdf>
- [16] A. Chaintreau, P. Hui, C. Diot, R. Gass, and J. Scott, “Impact of human mobility on the design of opportunistic forwarding algorithms,” in *Proc. of IEEE INFOCOM’06*, Barcelona, Spain, Apr. 2006.
- [17] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic, “Power law and exponential decay of inter contact times between mobile devices,” in *Proc. of ACM MobiCom’07*, Montreal, Canada, Sep. 2007.
- [18] P. Hui, J. Crowcroft, and E. Yoneki, “Distributed community detection in delay tolerant networks,” in *Proc. of SIGCOMM Workshop MobiArch’07*, Kyoto, Japan, Aug. 2007.
- [19] V. Erramilli, A. Chaintreau, M. Crovella, and C. Diot, “Diversity of forwarding paths in pocket switched networks,” in *Proc. of ACM IMC’07*, San Diego, CA, USA, Oct. 2007.