HERO – A Home Based Routing in Pocket Switched Networks*

Shengling Wang[1], Min Liu[1], Xiuzhen Cheng[2,4], Zhongcheng Li[1], Jianhui Huang[3], and Biao Chen[4]

[1] Institute of Computing Technology, Chinese Academy of Sciences, China
[2] Department of Computer Science, The George Washington University, USA
[3] IBM China
[4] Department of Computer Information Science, University of Macau, China

Abstract. Pocket switched networks (PSNs) take advantage of human mobility to distribute data. Investigations on real-world trace data indicate that human mobility follows a simple reproducible pattern: a human being usually visits a few places at high frequencies. These most frequently visited places form the *h*ome of a node, which is exploited in this paper to design two HomE based ROuting (HERO) algorithms. In the basic HERO, the first encountered relay whose home contains the place where the destination resides is selected to deliver the data. The enhanced HERO, on the other hand, continuously selects a better relay that visits the destination place at a higher frequency. In both algorithms, each node only needs to maintain and exchange its relatively stable home information and/or the corresponding visiting frequencies; therefore no global networking information and no frequent information update are needed, resulting in a low burden on the network due to its low communication and storage overheads. Moreover, HERO involves only simple arithmetic operations, thus causing little computation overhead at the mobile nodes. The simulation results indicate that both HERO algorithms outperform the state-of-the art.

Keywords: Pocket switched networks, routing, Human mobility.

1 Introduction

PSN is a new networking paradigm that makes use of human mobility to provide occasional contact opportunities for mobile devices to deliver data. It falls into the category of delay/disruption tolerant networks (DTNs). PSNs inherit some traits from DTNs such as intermittent connectivity, limited network capacity, and energy and storage constraints of the participants. These traits require PSN

* This work has been supported by the National Basic Research Program of China (No. 2011CB302702, No.2011CB302800), the Beijing Nova Program, the National Natural Science Foundation of China (No. 61120106008, No. 61132001, No. 61133015, No. 60803140, No. 60970133, No. 61070187, and No. 61003225), the National Science Foundation of the US (CNS-0831852), and the NPST program by King Saud University Project No. 10-INF1184-02.

X. Wang et al. (Eds.): WASA 2012, LNCS 7405, pp. 20–30, 2012.

applications to be delay/disruption tolerant and make *store-carry-forward* the mainstream communication mode for data delivery. In a typical PSN, mobile nodes[1] serve as relays to physically carry the data and forward it opportunistically upon contacting with others. Hence, relay selection is a key problem, which directly affects the efficiency of PSN data delivery.

Whether a node is a good relay in a PSN depends on the connectivity between the node and the destination. Nevertheless, it is a challenge to measure such a connectivity because node mobility makes it nondeterministic and dynamic. Moreover, relay selectors in a PSN are regular nodes with limited resources (storage, computation capability, battery power, etc.), and hence the burden of a selector for computing or comparing the connectivity with the nodes it contacts increases with the rapidly growing number of portable devices in the network. Therefore an effective and efficient routing strategy should place a low burden on the nodes and should be scalable. To achieve this goal, the routing strategy should satisfy the following three design requirements: i) no global knowledge maintained at each node; ii) no frequent information update at each node; and iii) low computation and storage overheads on each node. As indicated in Section 2, none of the existing routing strategies designed for PSNs could simultaneously satisfy all these three requirements.

In fact, the traits of PSNs are rooted from human mobility, which can simplify the routing decision. Our investigation [1] on the real world data reveals that human mobility has an important characteristic, namely, *a high degree of spatial regularity*. To be specific, each node has a significant probability of returning to a few highly frequently visited places. Taking into account the spatial regularity of human mobility, we propose two HomE based ROuting (HERO) algorithms for PSNs, with each placing at most one copy of a data in the network at any instant of time. Our algorithms introduce a concept of *home*, which is a set of places a node often visits. Because the probability that a node comes back home is high, the node can successfully deliver the data to the destination at a high probability if its home includes the place where the destination resides.

The basic HERO relies on the first encountered relay whose home contains the place where the destination resides to deliver the data. The enhanced HERO continuously changes the relay if a new one with a higher visiting frequency to the destination place is met. These two HERO algorithms have the following characteristics, which demonstrate that they do scale well, and do cause low burden on the network.

1) HERO requires no global networking knowledge maintained at each node, which results in low storage and communication overheads. Existing routing schemes such as [2–5] force each node maintain an entry for every other node in the network, while [6, 7] require each node to keep the global routing information and update it whenever the network topology changes. Compared to them, HERO orchestrates the spatial regularity of human mobility to determine the relay based on the local information, i.e, the most frequently visited places, which incurs little overhead.

---

[1] In our paper, a node refers to a human being.

2) HERO requires no frequent information update. In HERO, each node maintains the list of the most frequently visited places, which is relatively stable in a dynamic network environment. Moreover, this maintained information is relatively integral and precise, which results in infrequent updates. In some DTN routing algorithms such as those proposed in [6, 7], each node infers the global routing table (made up of a series of relays) from its local observations. Due to the locality constraint of each node, the global routing table may not be integral, correct, or consistent. As a result, the nodes may respond differently or even improperly.
3) HERO owns the distributed trait. Many existing PSN routing schemes such as those proposed in [8–10] employ a central node to bridge different communities, which may become the potential bottleneck. In HERO, no central role is involved. Hence, HERO is distributed in nature.
4) HERO is simple and effective. HERO involves only simple algebraic operations, resulting in low computation overhead. Moreover, HERO does not rely on any complicated or unrealistic human mobility model. For example, it does not require nodes to have strict repetitive motions such as in [11] or require the inter-contact time between two nodes to follow a specific distribution. Our simulation study based on the Dartmouth College mobility trace data validates its effectiveness and practicability.

It is worth noting that though the basic and enhanced HERO algorithms are presented as single-copy data delivery techniques for PSNs, their relay selection rules can be generalized naturally to get multi-copy versions. In this paper, we investigate the relay selection rules for single-copy data delivery because we want to clearly demonstrate how *home* can help to improve the efficiency of data delivery without relying on the redundancies caused by multiple copy data delivery.

The rest of the paper is organized as follows. Section 2 presents the related work while Section 3 elaborates on our HERO model. The basic and enhanced HERO algorithms are detailed in Section 4. Our conclusions are presented in Section 6.

2 Related Work

PSNs fall into the DTN category. Hence we first briefly review the major DTN routing algorithms in this section. Following that we address the popular routing algorithms designed for PSNs.

Existing DTN routing algorithms are classified into two categories: *deterministic* and *stochastic*. Deterministic approaches [12, 11, 13, 14] provide deterministic routing decisions assuming that some kinds of network connectivity information are known *a priori*. Jain *et al.* [12] modify Dijkstra's algorithm to compute the DTN routes (made up of a series of relays) when the network connectivity patterns are known. DHR [11] is a hierarchical routing framework based on the assumption that nodes in a network are either static or with strict repetitive motions. Conan *et al.* [13] minimize the delivery time given that the inter-contact interval between every pair of nodes is known. Gao *et al.* [14] formulate the problem of routing for multicast in DTNs as a unified knapsack problem assuming that the contact rate between any two nodes in the network is given.

Due to the uncertainty and dynamism of a DTN, it is challenging to obtain the network connectivity information. Hence, deterministic approaches are hard to implement in practical applications. This stirs the research of stochastic approaches [2, 3, 5, 15, 7, 16, 10].

PROPHET [2] uses the past encounters to predict the delivery probability. In FRESH [3], a node needs to keep a record of its most recent time meeting with each of the other nodes. Any node that encounters the destination more recently than the source can be selected as a relay. Gao *et al.* [5] exploit the transit contact pattern for each node, through which a node with a higher contact chance is selected as a relay. Liu and Wu [15] model the network as a probabilistic time-space graph and propose an expected minimum delay algorithm.

MaxProp [7] determines which data is transmitted or deleted from the buffer according to the delivery likelihood, whose computation requires each node i to keep track of f_j^i, the probability of the next meeting node being j. Dang and Wu [10] propose a cluster-based routing algorithm, in which nodes within the same cluster communicate directly while two nodes belonging to different clusters utilize gateways to relay the data.

Epidemic [16] selects relays randomly. It disseminates a large number of copies of each data in order to enhance the delivery ratio, which incurs a heavy communication overhead. To trade off between the communication overhead and the delivery ratio, a utility-based spraying method is proposed in [4], which requires each node i to maintain a utility function $U_i(j)$ for every other node j in the network and selects relays according to the utilities of the nodes.

Because PSNs are formed by human beings, their data delivery efficiency can be greatly enhanced by taking advantage of the traits of human behaviors. According to the traits of human behaviors an algorithm employs, social-based and location-based approaches are proposed, with the former making use of the sociality of human beings while the latter utilizing the spatial characteristics of human mobility to select relays.

BUBBLE [8], SimBet[9], SocialCast [17], user-centric dissemination [18], and SANE [19], are social-based mechanisms, where a more popular person has a higher chance to be utilized as a relay. MobySpace [20] is a location-based approach. It selects a node with a similar mobility pattern to the destination as the relay. The mobility pattern of a node is characterized by its probabilities of visiting all locations in the network.

3 The HERO Model

In this section, we introduce the HERO model in detail.

HERO divides the whole PSN area, denoted by Ω, into multiple zones Z_i, with $\cup Z_i = \Omega$ and $\cap Z_i = \emptyset$. These zones can have any shape. Each zone Z_i is identified by its center coordinates (x_i, y_i).

HERO can employ various methods for its nodes to figure out the center coordinates of a zone. For example, the center coordinates of a zone can be broadcasted by the access points or access routers in an infrastructure-based network; or they can be determined based on a mapping function if the node is aware of its own physical location.

As we have articulated earlier, human mobility follows a simple reproducible pattern: *one usually visits one or a few zones at a high frequency*. These frequently visited zones form the *home* of a node:

Definition 1 (Home). *The home of a node i, denoted by H_i, is the set of zones it usually visits.*

Home is the base of HERO. Thus it is critical to determine the home for each node. There exist two simple strategies: i) a node can statically configure the zones it usually visits as its home; and/or ii) it can dynamically add a zone to its home once the visiting frequency of the zone is larger than a given threshold. Similarly, a zone can be deleted from a node's home either statically or dynamically.

We assume that any two nodes located at the same zone can communicate directly with each other. In our HERO algorithms, once two nodes contact, they exchange their home information. A relay can be selected according to the distances between the zones of its home and that of the destination. Some related definitions are given as follows:

Definition 2 (Neighbor set). *The* neighbor set *of node i, denoted by N_i, is the set of nodes that can communicate directly with i.*

Based on our assumption, all the nodes covered by the zone where i resides belong to N_i. More generally, N_i includes the nodes in a neighboring zone that can communicate with i directly. Note that $i \notin N_i$.

Definition 3 (Destination zone). *The zone the destination currently resides is the* destination zone*, denoted by Z_d.*

Definition 4 (Distance between home and destination). *The distance between the home of node i, H_i, and the destination zone Z_d is the minimum distance between Z_d and any zone in H_i, i.e., $\|H_i - Z_d\| = \min\{\|Z_i - Z_d\| \mid Z_i \in H_i\}$.*

Definition 5 (Home node). *Node i is called a* home node *of a data with destination zone Z_d if $\|H_i - Z_d\| = 0$.*

4 HERO Algorithms

In this section, we propose the basic and enhanced HERO algorithms.

4.1 The Basic HERO Algorithm

When the source cannot directly communicate with the destination, the basic HERO algorithm selects a home node of the data as a relay. This relay is the only one for the data: once the relay receives the data, it never delivers to other nodes except the destination. The process of the basic HERO is given in Algorithm 1, where the function $send(B, i)$ indicates that the data B is sent to node i.

Algorithm 1. The Basic HERO Algorithm

Require: N_s: the neighbor set of the source s; Z_d: the destination zone; H_i: the home of node i.

repeat
  Update N_s
  ▷ If destination d is in the neighborhood, deliver directly;
  for each node $i \in N_s$ **do**
    if $i = d$ **then**
      $send(B, i)$, **return**
    end if
  end for
  ▷ If the source is a relay, no other relay node needs to be selected;
  if $\|H_s - Z_d\| = 0$ **then**
   return
  end if
  ▷ If locating a home node of the data in the neighborhood, selects this node as a relay;
  if $\exists i \in N_s, s.t. \|H_i - Z_d\| = 0$ **then**
    $send(B, i)$, **return**
  end if
until the data expires

4.2 The Enhanced HERO Algorithm

The basic HERO is simple and naive. Based on it, many variants can be produced to enhance the efficiency of data delivery. In this subsection, we elaborate an enhanced HERO algorithm that takes into account the visiting frequency of a node to a zone. As the visiting frequencies to the zones in a home are different, we introduce the concept of *visiting intensity* to depict this trait:

Definition 6 (Visiting intensity). *The* visiting intensity *of zone Z_j by node i, denoted as V_{ij}, is the visiting frequency of node i to zone Z_j within a unit time.*

The enhanced HERO is a single-copy multi-relay mechanism based on the concept of home. In this algorithm, once two nodes contact, they exchange their home and visiting intensity information. The main idea of the enhanced HERO is to continuously find the relay whose probability to visit the destination zone is higher than that of the node currently carrying the data. When a source cannot communicate with the destination directly, it delivers the data to a home node in its neighborhood whose visiting intensity is the highest. If this node meets another home node whose visiting intensity is higher than other neighboring home nodes and itself, it delivers the data to this home node and discards the data itself. The detailed description of the enhanced HERO is shown in Algorithm 2.

Algorithm 2. The Enhanced HERO Algorithm

Require: N_i: the neighbor set of node i; Z_d: the destination zone; S_{Hi}: the set of home nodes met by the node i; V_{ij}: the visiting intensity of node i to zone Z_j.

repeat
  Update N_i ▷ Node i may be the source or a relay;
  for each node $j \in N_i$ **do**
    if $j = d$ **then**
      $send(B, j)$, **return**
    end if
    Update N_j
  end for
  $S_{Hi} \leftarrow \{j \in N_i \mid \|H_j - Z_d\| = 0\}$
  if $S_{Hi} \neq \emptyset$ **then**
    ▷ Select the neighbor k with the highest visiting intensity to Z_d;
    if $\exists k \in S_{Hi} \wedge (V_{kZ_d} > V_{iZ_d}) \wedge (V_{kZ_d} \geq V_{hZ_d}$ for $\forall h \in S_{Hi} \backslash \{k\})$ **then**
      $send(B, k)$, **return**
    end if
  end if
until the data expires

Note that both the basic HERO and the enhanced HERO are inherently based on the following assumption: the destination stays at Z_d during the process of data delivery. This assumption is reasonable especially when nodes connected through WiFi, the most mainstream wireless technology used in PSNs, because many real world trace records such as [21, 22] indicate that the mobility pattern of WiFi nodes is quasi-static in a sense that the clients tend to stay in the same location for a long time. However, if this assumption does not hold, our algorithms can still work by slightly changing their ways of usage. For example, the data can be delivered to each zone in the destination's home in light of the routing policy of our algorithms. Even though the destination is not located at any zone of its home, the data can still be delivered to the infrastructure or a static node in the zones of the destination's home, from which the data can be retrieved when the destination comes back home.

5 Performance Validation

In this section, we evaluate the performance of HERO with the Dartmouth College mobility trace data [23]. We choose the data collected from 09/21/2003 to 10/20/2003 because in this period the records are integral and the nodes' behaviors are regular.

In our simulation study, each AP is represented as a zone. As described above, by changing the ways of utilizing our algorithms, the data can be delivered to a destination with high mobility. Hence, in this simulation, we keep the basic assumption that the destination stays at Z_d during the process of data delivery. We randomly choose 100 mobile nodes as the sources and randomly assign one of the APs to each source as its destination. Thus, there are in total 100 communication pairs, which remains unchanged in our simulation study.

Because the running time increases rapidly as the number of mobile nodes increases, we limit the number of nodes to a manageable size, a common measure taken by [6, 20, 4], which also use the Dartmouth College trace data. To construct the simulation scenarios, we first randomly select 200 mobile nodes and add the 100 sources selected before to get a 300-node scenario. Then we add 100 randomly selected new mobile nodes to the 300-node scenario to get a 400-node scenario. Repeat this process we obtain the 500-node and 600-node scenarios. For each network scenario, we repeat 10 times and the averaged results are reported to enhance the confidence level. For simplicity, we denote by $U = x$ the x-node scenario, where U is the network size. Because 300˜600 nodes represent about 5.4˜10.8% of the total mobile nodes in the trace, which contains 5543 mobile nodes, the performance is worse than that obtained from the whole set of trace date for each scheme investigated in this simulation study.

Let the number of zones in a home be the home size. Through extensive tests we found that when the largest home size is limited to 10% of the zones in the network, namely the average home size is about 10, a good trade-off between quality and quantity of relays can be obtained.

In this simulation, we compare the performance of HERO with that of MobySpace [20], the most related research that selects a relay with a similar mobility pattern as the destination, and that of Epidemic [16], a flooding algorithm

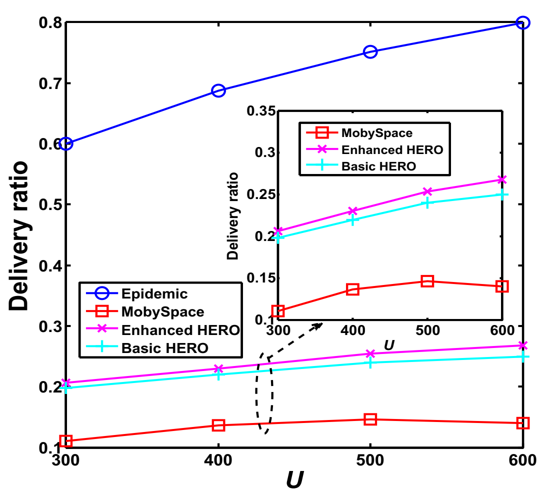

Fig. 1. Comparison of delivery ratio

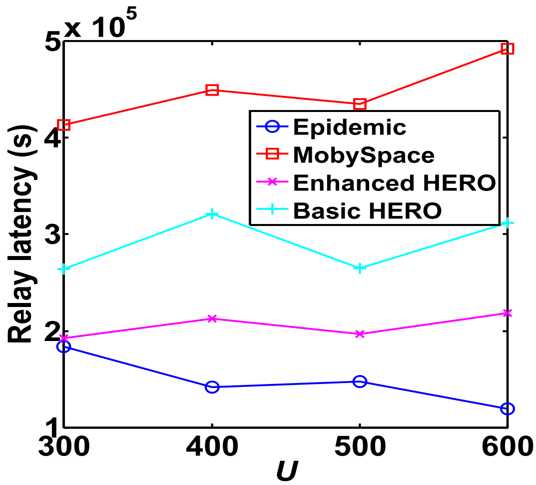

Fig. 2. Comparison of relay latency

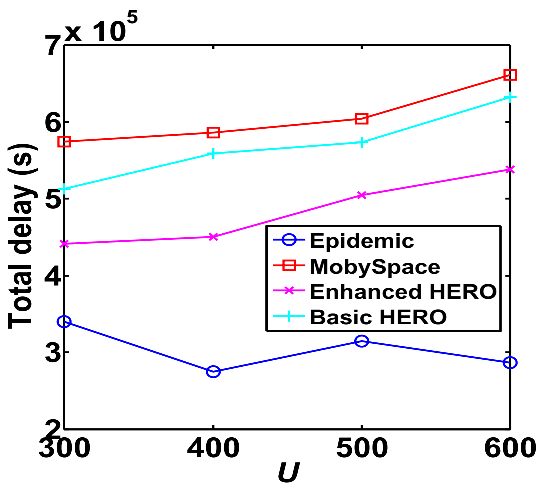

Fig. 3. Comparison of total latency

serving as the base for comparison study. The data lifetime equals simulation duration.

Figs. 1, 2 and 3 report the delivery ratio, the average relay latency, and the average total latency of the four schemes, respectively. Because HERO and Mobyspace are all single-copy algorithms, we use the axes graph in Fig. 1 to emphasize their difference.

Fig. 1 indicates that the data delivery ratio of Epidemic is the highest and that of MobySpace is the lowest. In addition, the data delivery ratio of all algorithms except Epidemic is slightly increased when the network size U increases. These observations can be justified as follows. Since Epidemic adopts a flooding policy, it can make a better use of the network connectivity, which is enhanced significantly when U increases. However, in HERO and Mobyspace, no matter how many mobile nodes exist, there is at most one copy of the to-be-delivered data in the network at any instant of time. As a result, though the probability of finding a good relay is increased when U increases, the enhancement of the network connectivity only slightly impacts on the performance of HERO and Mobyspace. This is the reason why the delivery ratio of Epidemic is increased faster than those of HERO and Mobyspace.

From Figs. 2 and 3, we observe that the average relay latency and the average total latency of Epidemic are the shortest, while those of Mobyspace are the longest and those of HERO algorithms are in-between. In addition, their average relay latencies and average total latencies are susceptible to the relay latencies and total latencies of the newly added successful communication pairs when U increases. Hence, when U increases, their average relay latencies and average total latencies either increase or decrease.

Based on the above analysis, we conclude that the performance of Epidemic in terms of the data delivery ratio, relay latency, and total latency is better than that of the enhanced HERO, which is better than that of the basic HERO. The performance of MobySpace is the worst.

6 Conclusion

In this paper, we propose two home-based routing (HERO) algorithms, the basic HERO and the enhanced HERO, which make use of the spatial regularity of human mobility to select relays. The basic HERO is a single-copy single-relay algorithm while the enhanced one is a single-copy multi-relay algorithm. Both algorithms rely on the concept of *home*, which is the set of places a node often visits. We use the Dartmouth college trace data to validate the performance of both HERO algorithms in terms of data delivery ratio, relay latency, and end-to-end delay, and compare them with two relevant research, MobySpace and Epidemic. Our simulation results indicate that both HERO algorithms outperform Mobyspace but are worse than Epidemic, which provides an upper bound on the delivery ratio and a lower bound on the delivery latency. However, the transmission cost of Epidemic in terms of the number of relays is much higher than those of the HERO algorithms.

References

1. Wang, S., Liu, M., Cheng, X., Song, M.: Routing in Pocket Switched Networks. IEEE Wireless Communications 19(2), 67–73 (2012)
2. Lindgren, A., Doria, A., Scheln, O.: Probabilistic routing in intermittently connected networks. In: MobiHoc, Annapolis Maryland, USA (2003)
3. Dubois-Ferriere, H., Grossglauser, M., Vetterli, M.: Age matters: efficient route discovery in mobile Ad hoc networks using encounter ages. In: MobiHoc, Annapolis, Maryland, USA (2003)

4. Spyropoulos, T., Turletti, T., Obraczka, K.: Routing in Delay-Tolerant Networks comprising heterogeneous node populations. IEEE Transactions on Mobile Computing 8(8), 1132–1147 (2009)
5. Gao, W., Cao, G.: On exploiting transient contact patterns for data forwarding in delay tolerant networks. In: ICNP, Kyotp, Japan (2010)
6. Jones, E., Li, L., Schmidtke, J., Ward, P.: Practical Routing in delay tolerant networks. IEEE Transactions on Mobile Computing 6(8), 943–959 (2007)
7. Burgess, J., Gallagher, B., Jensen, D., Levine, B.: MaxProp: routing for vehicle-based disruption-tolerant networks. In: INFOCOM, Piscataway, USA (2006)
8. Hui, P., Crowcroft, J., Yoneki, E.: BUBBBLE Rap: social-based forwarding in delay tolerant networks. In: MOBIHOC, Hong Kong, China (2008)
9. Daly, E., Haahr, M.: Social network analysis for routing in disconnected delay-tolerant MANETs. In: MOBIHOC, Montreal, CA (2007)
10. Dang, H., Wu, H.: Practical Clustering and cluster-based routing protocol for delay-tolerant mobile networks. IEEE Transactions on Wireless Communication 9(6), 1874–1881 (2010)
11. Liu, C., Wu, J.: Scalable routing in delay tolerant networks. In: MOBIHOC, Montreal, CA (2007)
12. Jain, S., Fall, K., Patra, R.: Routing in a delay tolerant network. In: SIGCOMM, New York, USA (2004)
13. Conan, V., Leguay, J., Friedman, T.: Fixed point opportunistic routing in delay tolerant networks. IEEE Journal on Selected Areas in Communications 26(5), 773–781 (2008)
14. Gao, W., Li, Q., Zhao, B., Cao, G.: Multicasting in delay tolerant networks: a social network perspective. In: MOBIHOC, New Orleans, USA (2009)
15. Liu, C., Wu, J.: Routing in a cyclic Mobispace. In: MOBIHOC, Hong Kong, China (2008)
16. Vahdat, A., Becker, D.: Epidemic routing for partially connected ad hoc networks. Technical Report CS-200006, Duke University (2000)
17. Costa, P., Mascolo, C., Musolesi, M., Picco, G.: Socially-aware routing for publish-subscribe in delay-tolerant mobile Ad hoc networks. IEEE Journal on Selected Areas in Communications 26(5), 748–760 (2008)
18. Gao, W., Cao, G.: User-centric data dissenmination in disruption tolerant networks. In: INFOCOM, Shanghai, China (2011)
19. Alessandro Mei, A., Morabito, G., Santi, P., Stefa, J.: Social-aware stateless forwarding in pocket switched networks. In: INFOCOM, Shanghai, China (2011)
20. Leguay, J., Friedman, T., Conan, V.: Evaluating mobility pattern space routing for DTNs. In: INFOCOM, Barcelona, Catalunya, SPAIN (2006)
21. Balachandran, A., Voelker, M., Bahl, P., Rangan, P.: Characterizing user behavior and network performance in a public wireless LAN. In: ACM SIGMETRICS International Conference on Measurement and Modeling of Computer Systems, Marina Del Rey, California (2002)
22. Balazinska, M., Castro, P.: Characterizing mobility and network usage in a corporate wireless local-area network. In: MobiSys 2003, San Francisco, California (2003)
23. Henderson, T., Kotz, D., Abyzov, I., Yeo, J.: CRAWDAD trace set dartmouth/campus/movement (v.2005-03-08) (2005), http://crawdad.cs.dartmouth.edu/dartmouth/campus/movement