Empirical Evaluation of Hybrid Opportunistic Networks

Pan Hui*†, Anders Lindgren*, Jon Crowcroft*
firstname.lastname@cl.cam.ac.uk
* University of Cambridge
† Deutsche Telekom Laboratories

Abstract— **In this paper, we study the utility of opportunistic communication systems with the co-existence of network infrastructure using three real experimental deployments. We study how some important performance metrics change with varying degrees of infrastructure and mobile nodes willing to participate in the opportunistic forwarding. In doing so, we observe phase transitions in the utility of infrastructure and opportunistic forwarding respectively at different points in the design space. We discuss the implications that this has for the design of future network deployments and how this observation can be used to improve network performance, while keeping cost at a minimum. To the best of our knowledge, this is the largest scale of experimental evaluations of hybrid opportunistic networks. It provides empirical evaluations to the large body of theoretical work that has been done in the past and insights for future research.**

I. INTRODUCTION

We can observe two trends in the development of communication networks. One trend is that as mobile devices have become more and more popular, the possibilities of deploying mobile ad hoc and opportunistic networks have increased. It is more and more likely that communication relying on peer-to-peer local connectivity can become a reality. This is an active research topic in the area of mobile ad hoc networking and delay tolerant networking [8] [10] [11]. On the other hand, with the decrease in the price of wireless routers and Internet access service, it is also possible that wireless access networks can provide a nearly complete coverage of a whole city area [26]. There is therefore an abundance of sceptics to delay tolerant and opportunistic communication who claim that such modes of communication will be unnecessary, as one can rely solely on the presence of a network infrastructure. In this paper, we address these issues by investigating under which conditions opportunistic communication will be necessary or useful for successful network operation, and how much improvement it can provide over a purely infrastructure-based system. We also look at the converse issue of how the possible performance of opportunistic ad hoc systems can be improved by making use of whatever infrastructure (albeit sparse) is available.

We investigate these issues through the use of simulations using three sets of experimental measurement data. The first data set contain traces from a 3-day contact measurement experiment done during the Infocom 2006 conference. These traces contain data about contacts between both mobile users and stationary nodes spread over the conference site (representing an infrastructure). We also use traces from a 9-month contact experiment done at MIT with students and researchers as participants. These traces contain data about contacts between mobile users, as well as data about which cellular towers the mobile users are connected to. We use these two experiments to evaluate nonuniform deployments of infrastructures. To also evaluate a uniform deployment of access points (APs), we use a data set from the KAIST university campus in Korea, which contains Global Position System (GPS) locations of the mobile devices. We deploy virtual APs by assuming wireless IEEE 802.11 communication range. APs are deployed virtually on a grid in the area within which most of the node movements take place. Taking this to a larger scale, this can also show similarities to, for example, a well covered WiFi city area, and by changing the number of APs that we choose to use in our analysis, we are able to represent different densities of infrastructure coverage.

To study the system, we focus on two different classes of applications that are likely to be prevalent and popular in a system like this. The first application is an asynchronous peer-to-peer unicast messaging system (e.g., a form of mobile IM, which in a recent real-life DTN deployment was one of the most popular applications among users [19]), in which mobile users will send asynchronous messages directed to other mobile users. The second application is a data push service, such as an email or news delivery service, in which messages are injected into the mobile network at the access points and are targeted to the mobile nodes. We study the performance of such systems at different access point densities, and then proceed to investigate the possible benefits of using opportunistic peer-to-peer communication between the mobile nodes to improve system performance. We choose to evaluate the performance in terms of delivery ratio and end to end delay. While metrics such as system cost, resource utilisation, and power consumption are be important, they are not in the scope of this paper.

We observe that opportunistic communication has a significant impact on improving the throughput and also average delivery delay of both asynchronous messaging and the data push service, especially when the density of the infrastructure is low. We define a metric to measure the utility of opportunistic network in the coexistence of infrastructure and vice versa for the infrastructure. We observe a phase transition [14] for the utility of opportunistic communication with the change of infrastructure density. This indicates that as the number of active access points in the system reaches a certain point, the utility of opportunistic communication remains at a similar level for further increases in infrastructure density. Thus, we show that even with a fairly large infrastructure presence, there is a significant utility in using opportunistic communication. We also observe the phase transition for the infrastructure with the coexistence of opportunistic network.

Well-planned infrastructured networks are in general efficient in providing low-latency delivery of data messages, while opportunistic networking provides a viable alternative if applications are more tolerant to delay. Installing wired base stations can significantly lower delays and increase delivery; but the costs can be as high as US$5,000 per base station [15] and the number of APs necessary to produce significant improvement is also difficult to determine. In one of the experiments, we show that utilizing opportunistic networking is more efficient than uniformly deploying 900 APs in the area where most of the node activity is happening, which would cost as much as US$4,500,000. Knowing when the use of opportunistic communication will yield satisfactory performance can save a lot of money by not deploying infrastructure. At the same time, it is also important for deployment planning to know the situations where infrastructure is necessary to guarantee certain delivery and delay bounds. This paper serves the purpose of providing insights and simple guidelines on deciding so by empirical evaluations of three experimental system. We are not targeting a theoretical bound for the analysis, but draw general conclusions from empirical observations, which can be useful for future theoretical works to step on. Our previous work [12] showed some preliminary work using one dataset, the contribution of this paper is to generalise the observations from several experiments and study infrastructure deployments on a larger scale.

In the following section we describe the application scenarios in more detail. In Section III, the methodology used in our analysis is described and the results from it are shown in Section IV. Some related work is discussed in Section V, and finally Section VI discusses the implications of our findings and concludes.

II. APPLICATION SCENARIO

To evaluate the performance of the system, we look at two application scenarios that are likely to be common in the target networks. The applications we use for our study are an *asynchronous messaging* service and a *data push* service. For both systems, we assume that the access points are all connected to each other through an underlying high speed network, and that two nodes that are simultaneously connected to active access points are able to communicate with each other without additional delay.

- **Asynchronous Messaging.** In this application scenario, mobile nodes wish to exchange messages between each other. This is likely to happen in applications where users want to send messages directed directly to another users in a service similar to instant messaging or SMS. This was recently shown to be a very popular application in real-life DTN deployments[19].
  There are three possible modes of operation for this application. In the first mode of operation, the source of a message will only transmit it directly to the destination when the two nodes are in direct contact, or when in contact with each other through the infrastructure (*Only APs* in the results). The other alternative is for the nodes to use opportunistic collaborative forwarding in the network to propagate the messages (*Only Opportunistic Flooding* and *Only Opportunistic MCP* in the results). Finally, we allow for the two delivery options to be used in combination with each other, allowing nodes to both do opportunistic forwarding, as well as using the infrastructure to traverse the network.
- **Data Push.** In this application scenario, all mobile nodes subscribe to a data push service such as, for example, delivery of e-mail or a news feed from the access points. For this service, messages are being generated in the infrastructure and addressed to mobile nodes. In the basic, traditional, system, nodes can only access their messages as they encounter an access point themselves. In order to improve the performance of the system and reduce the delay experienced by the mobile nodes to receive messages from the access points, we allow messages to be picked up by other mobile nodes as well and opportunistically transmitted between the mobile nodes until reaching the destination.

III. METHODOLOGY

In order to evaluate the system and study the impact that varying degrees of infrastructure and opportunistic participation have on the performance of the applications outlined in the previous section, we ran simulations using three different sets of experimental trace data A more detailed description of these datasets can be found in the following section. As we are looking at results on the network and session layers, we do not include details about lower layer interactions in our evaluations. Those are assumed to be handled appropriately, and give rise to the contacts in the trace file.

A. Experiment Dataset

In order to perform our analysis, we use three experimental data sets gathered by the Haggle project at the Infocom 2006 conference [12], the MIT Reality Mining project at the MIT campus [7], and the NCSU MobilityDTN project at the Kaist campus [25], known as *Infocom06*, *Reality*, and *Kaist* respectively. Additional technical information about the data sets can be found in Table I.

- In *Infocom06*, devices logging contacts with other Bluetooth devices through the device discovery protocol of Bluetooth were distributed to 80 conference participants (due to hardware problems, data could only be retrieved from 78 of those nodes). The participants were asked to carry these contact loggers for the duration of the conference. In addition, 20 devices with external antennas

Experimental data set	Infocom06	Reality	Kaist
Device	iMote	Phone	Phone
Network type	Bluetooth	Bluetooth	GPS
Duration (days)	3	21	0.33
Granularity (seconds)	600	300	10
Number of experimental devices	98	97	92
Number of internal contacts	191,336	11,962	40,218
Average no. of contacts/pair/day	6.7	0.061	14.412

TABLE I

CHARACTERISTICS OF THE THREE EXPERIMENTAL DATA SETS

(providing longer range) were deployed at several places at the conference venue to act as access points. The placement of the access points (AP) are shown in Figure 1.

- In *Reality*, 100 smart phones were deployed to students and staff at MIT over a period of 9 months (we select a session of three-week period for our simulations). These phones were running software that logged contacts with other Bluetooth enabled devices by doing Bluetooth device discovery every five minutes, as well as logging information about the cellular tower they are associated with (a total of 31545 different towers were logged). In this paper, we use the cellular towers to mimic the wireless APs. Since the cellular networks usually provide full coverage in an urban area, which can not be expected for the types of infrastructures we envision here, we chose to only use a subset of the cellular towers as access points. We selected the 253 towers that had been most accessed by the mobile nodes (these were the cellular towers that have at least 100 contacts with mobile nodes logged) to represent the APs.
- In *Kaist*, Garmin GPS 60CSx handheld receivers were used for data collection which are WAAS (Wide Area Augmentation System) capable with a position accuracy of better than three meters for 95 percent of the time. The GPS receivers take readings of their current position every 10 seconds and record them into a daily track log. There are in total 92 traces from this dataset. There is no physical deployment of wireless APs in this experiment, but since we have the GPS locations of the devices, we can deploy virtual APs in the environment. In this paper, we select the area with most of the mobilities and deploy virtual APs with varying densities (the area covered by the coordinates (-5000, -5000) to (5000, 5000) in Figure 2). We assume the communication range for the mobile devices and the APs to be 100 meters.

B. Simulation Evaluation

For the simulation evaluation, we developed a simulator called *HaggleSim*, which can replay the collected mobility traces, and simulate different forwarding strategies on every contact event. This simulator is completely driven by contact events. The original trace file is divided into discrete sequential contact events, and fed into the simulator as input. In all the simulations in this work, we divided the trace into discrete contact events with a granularity of 100 seconds.

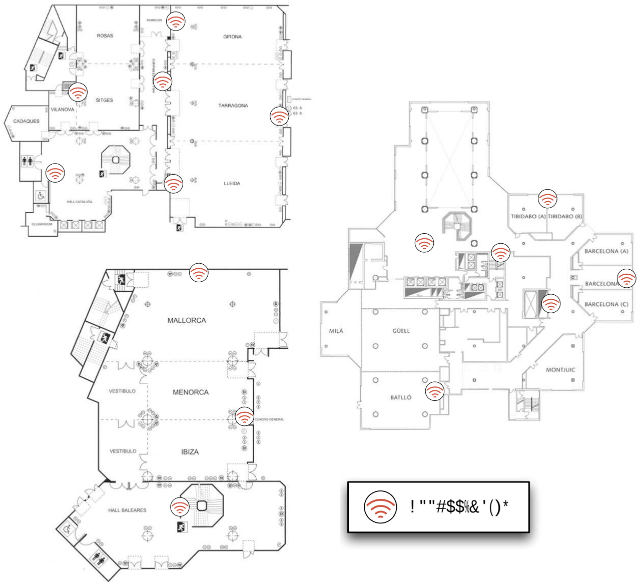


Fig. 1. Access point placement. In addition to the access points shown on this map, one was deployed with the hotel concierge, one at the hotel bar, and three in the lifts of the hotel.

We use the simulator to simulate a number of different forwarding strategies to see how they compare. The main forwarding strategies evaluated are:

- Opportunistic Flooding. This is a pure opportunistic flooding strategy in which messages are duplicated and sent to all mobile nodes encountered, as in epidemic routing[29].
- Opportunistic MCP. This is the controlled flooding scheme introduced in [11]; Multiple copies are sent subject to a hop count limit on the propagation of messages.
- Only APs. This is a pure infrastructure scenario, in which messages are only passed through access points.

We also simulate the variations of Flooding and MCP where access points are present and can be used by the protocols.

We analyse the successful *delivery ratio*, which is the proportion of messages that have been delivered out of the total unique messages created. Messages in the system are given a time to live value after which they are no longer considered of interest. By varying this value, the delivery ratio also reflects the delay of the system.

In order to evaluate the system performance of opportunistic network and also the infrastructure, we define a metric called *Relative Network Utility* (we will henceforth refer to it as *Utility* in this paper), which is based on the concept of information centrality in complex network analysis [5]. Information centrality relates the node centrality to the ability of the network to respond to the deactivation of the node. The information centrality of node i is defined as the relative drop in the network efficiency $E[G]$ caused by the removal from G of the edges incident in i. The efficiency of a network is a quantity which measures how well the nodes of the

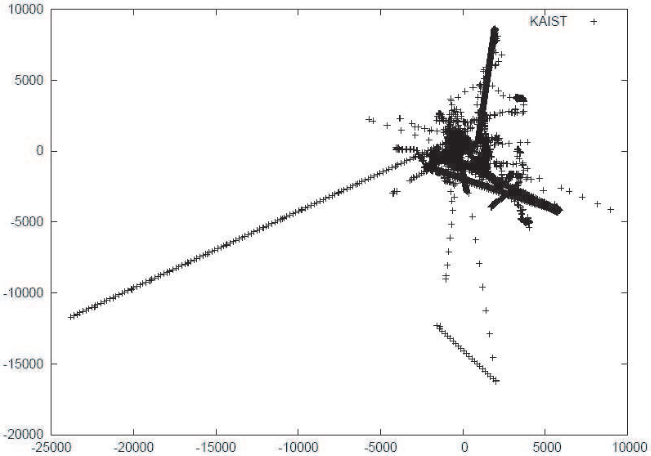


Fig. 2. The movement of all the nodes in the Kaist data set with anonymous GPS location information. In this paper, we select to place an infrastructure grid in the area covered by the coordinates (-5000, -5000) to (5000, 5000), which most of the mobility takes place.

graph exchange information; a more detailed definition can be found in [5]. In a system with a mix of opportunistic communication and infrastructure, we define the utility of opportunistic communication as the drop in throughput of the system, when we remove the opportunistic relays. That is, basically, the throughput of the system when both opportunistic communication and infrastructure are being used minus the throughput of the system if only infrastructure was present, divided by the total throughput as defined by Eq. 1.

$$U(O) = \frac{T(O+I) - T(I)}{T(O+I)} \tag{1}$$

$T(O+I)$ is the throughput (delivery ratio) of the system with both opportunistic communication and infrastructure, and $T(I)$ is the throughput of the system with only infrastructure. Vice versa, we define the utility of the infrastructure as $U(I)$:

$$U(I) = \frac{T(O+I) - T(O)}{T(O+I)} \tag{2}$$

where $T(O)$ is the throughput of the system with only opportunistic communication.

For all the simulations conducted in this paper, we compute the average and the 95% confidence intervals (using the t-distribution) of the throughput/delivery ratio of each forwarding scheme. To achieve statistical significance, we run each simulation 20 times with different random seeds.

IV. RESULTS AND EVALUATIONS

In this section, we present the simulation results on the trace for the asynchronous messaging and the data push applications. The results for the *Infocom06* dataset have been briefly introduced in our preliminary work [12], so in this paper we will focus more on presenting the results for the *Reality* and *Kaist* datasets, but still draw some general conclusion from the *Infocom06* dataset.

A. Asynchronous Messaging

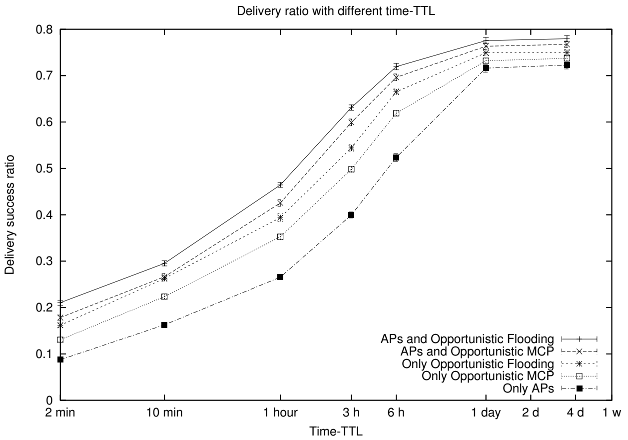
To evaluate the asynchronous messaging application, we create the following scenario: a total of 1000 messages are created in the system, with randomly selected sources and destinations; the message creation times are uniformly distributed throughout the experiment duration. Figure 3(a)(b) and (c) show the delivery ratio for asynchronous messaging with different forwarding schemes for the *Infocom06*, *Reality*, and *Kaist* datasets respectively. The x-axis shows the time-TTL, which is the maximum time each message can stay in the system.

We can see that the *APs and Opportunistic Flooding* scheme can double the throughput/delivery ratio of the *Only APs* scheme when the time-TTL is smaller than 1 hour, and give performance improvements of up to 63% and 40% when the time-TTLs are 3 hours and 6 hours in the *Infocom06* case. In the *Reality* case, deliveries within 1 hours happen mainly due to contacts with the APs. After that, the opportunistic schemes start to take effect and outperform the *Only APs* scheme when the time-TTL is set to 1 day. The performance of the *Only APs* scheme has no significant improvement after 1 day, but the delivery of the opportunistic schemes increase with time TTL until the end of the mobility traces.[1] Actually the deliveries by *Only APs* and *pure opportunistic* are nearly mutually exclusive in this experiment. The results when using both schemes combined are almost the direct sum of the deliveries of each scheme. For the *Kaist* case, we show the results when using pure opportunistic flooding, *Only APs* with the number of APs on each row of the grid set to 15, 30, and 50, and also the combined results of opportunistic flooding and 50 APs per row. When the density of APs is low (e.g. 15 APs per row), the deliveries caused by the APs are very low (0.15 with time TTL as high as 8 hours). We want to emphasise here that for this case, 15 APs actually means $15 \times 15 = 225$ APs in the whole system. When the number of APs increases to $30 \times 30 = 900$, the deliveries increase to a maximum of 70% with the maximum allowed time TTL, but this is still much lower than the opportunistic flooding case. When the APs density is as high as 2500 APs in the area, the deliveries are almost instant (2 minutes) and as high as above 70%. A combination of APs and opportunistic flooding provide further improvement for delivery, to as high as almost 100% totally delivery.

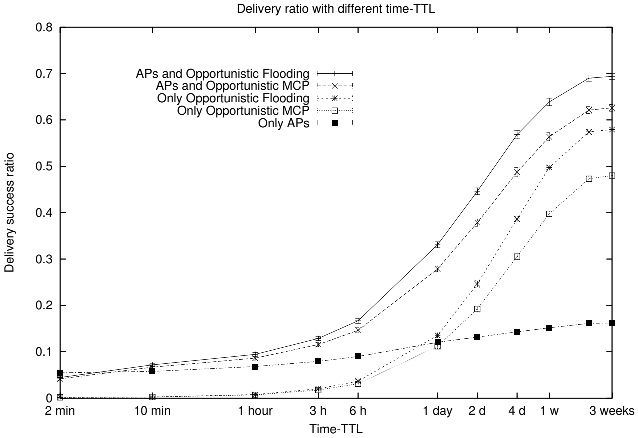
The delivery characteristics of these three experiments show some difference in terms of delivery ratio, time TTL required to achieve a particular delivery ratio, and the delivery ability of the APs, but a general conclusion we can draw is that opportunistic communication is very useful for improving the delivery ratio in all cases, especially in the *Reality* and *Kaist* datasets (relying solely on opportunistic communication yields

---

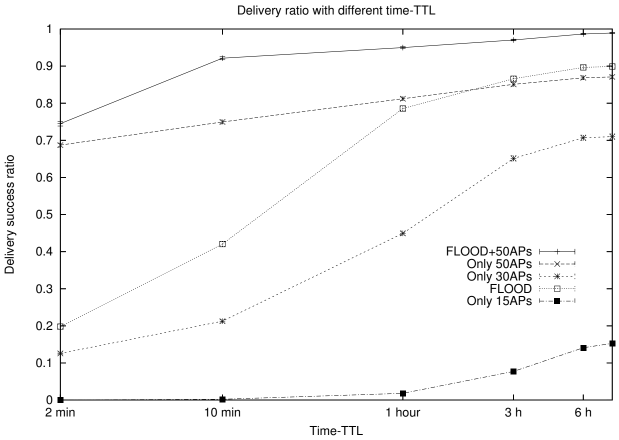
[1] The delivery ratio of the Reality dataset is generally low even with TTL up to one week. A possible reason is that many participants may switch off their Bluetooth transceivers to save batteries, which makes the network very sparse. But since the dataset lasts for 9 months (the longest available dataset) and has logging for cellular towers, can be a good subject for hybrid network study, we include it here.



(a) *Infocom06*



(b) *Reality*



(c) *Kaist*

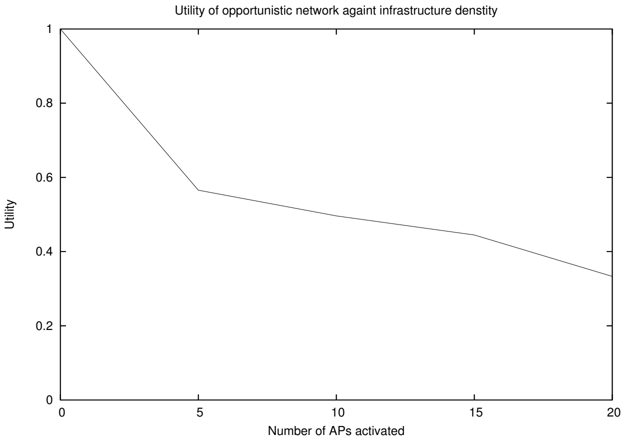
Fig. 3. Delivery ratio for asynchronous message delivery with time TTL

double the delivery ratio that one gets by deploying 900 APs in the Kaist system). We can see that there is a significant benefit of using opportunistic communication to improve the delivery ratio for asynchronous messaging in the system if the messages are not time critical. Figure 3(a)(b) and (c) gives an overall picture of the performance of opportunistic communication and infrastructure, but does not clearly quantify how useful each part of the system is with the coexistence of the other part. The network utility, U, that we introduced in Section III-B serves this purpose.

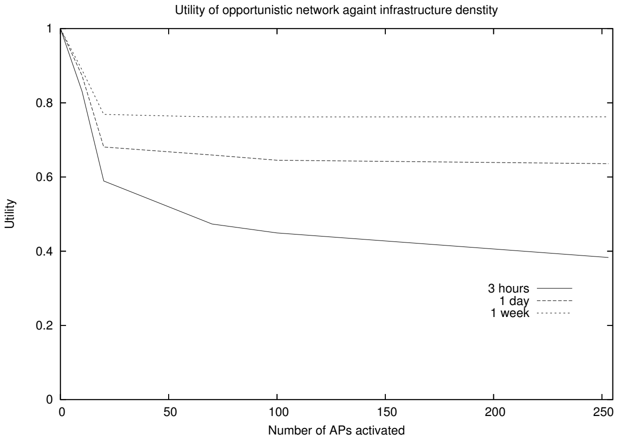
Figure 4(a) shows the utility of opportunistic communication with the coexistence of infrastructure for asynchronous messaging when the time-TTL is set to 3 hours. The x-axis shows the number of APs activated and the y-axis shows the utility of the opportunistic communication, $U(O)$. We observe a phase transition of the utility when the number of APs equals to 5, beyond that the utility decreases slowly with the increasing in number of APs and reaches 0.34 when the number of APs equals to 20. We can tell that the utility of the opportunistic communication decrease rapidly from 0 AP to 5 APs, which is the phase transition point. After the transition point, further increase in the number of APs does not have as large effect on the utility of the opportunistic communication. This means that when the number of APs in the system is small, the use of opportunistic communication is very important for the performance of the system in terms of throughput. When the number of APs is large, the utility of opportunistic communication decreases slowly as more access points are added and we can see that there is still a significant utility in using it when the number of APs reach 20. We can clearly say that opportunistic communication is important for the throughput of the system we are studying.

We further show the utilities of the opportunistic networks for the *Reality* and *Kaist* datasets in Figure 4(b) and Figure 4(c) with three different time TTLs. For *Reality*, we show the cases for time TTLs equal to 3hours, 1day, and 1 week. We see a deep drop of utility as we go from 0 to around 20 APs, and after that the utility decreases very slowly. The general trend for these three lines are very similar, except some difference in the values. In general, we can see that higher time TTL favors opportunistic communication, which is expected as it allows more time for the messages to be delivered, and opportunistic communication is well-tailored for delay-tolerant applications. When the time TTL is equal to 1 week, the utility stay as high as 0.8 no matter how many APs we include in the system. When the time TTL is set to 1 day, the utility is staying at minimum 0.62. Similar trend also observed in the *Kaist* dataset, but required more APs to reach the saturation (900 APs is the phase transition point). A different observation in this case is that the opportunistic network has higher utility when the time TTL is lower. The reason is because that in this experiment the opportunistic network is quite dense and hence about 40% of the deliveries can be done within 10 minutes by pure opportunistic transmission and at the same time the deliveries are quite low with low density of APs.

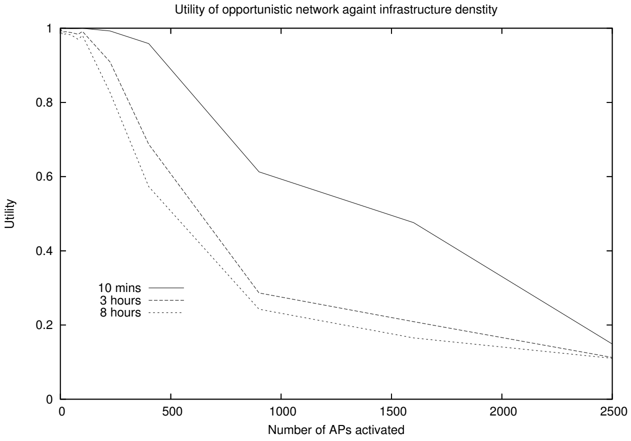
Figure 5(a) shows the utility of infrastructure with the



(a) *Infocom06*



(b) *Reality*



(c) *Kaist*

Fig. 4. Utility of opportunistic communication for asynchronous messaging with coexistence of infrastructure

coexistence of opportunistic communication. In this case, we use all the mobile nodes as relays and we also use the 3-hour time-TTL case. We observe a phase transition when the number of APs equal 15. Before that, the utility of the infrastructure is very low and increase very slowly. After the phase transition point, the utility increase rapidly but still only reaches 0.17 when the number of APs reaches 20, which shows that with the coexistence of the opportunistic communication, the utility of the infrastructure is quite low for improving the throughput of the system. Figure 5(b) shows the utility of the infrastructure for the *Reality* case. We observe a phase transition at around 20 APs; before that the utility is low but increases rapidly; after that the utility becomes stable with the increase in number of APs. The utility is generally higher for smaller time TTL, for example for the 3hours case, the utility reaches its maximum of above 0.8 when the number of active APs is over 100. Figure 5(c) further shows the utility for the *Kaist* case. We can see that the utility of the infrastructure is low (less than 0.2) for the cases of 3hour and 8hour time TTL. This is because the deliveries by the opportunistic network are quite high in this case and hence over weight the deliveries by the infrastructure. The utility can be up to 0.7 when the time TTL is set to 10 minutes. The phase transition again occurs around 900 APs (a grid with 30 APs on each row). The utilities of both infrastructure and opportunistic network (Fig. 5(c)) are high in the $10mins$ TTL case. This shows that in the *Kaist* case, both infrastructure and opportunistic network are important for providing short-delay deliveries. For long-delay-tolerant deliveries, any of the two systems can already provide satisfactory deliveries.

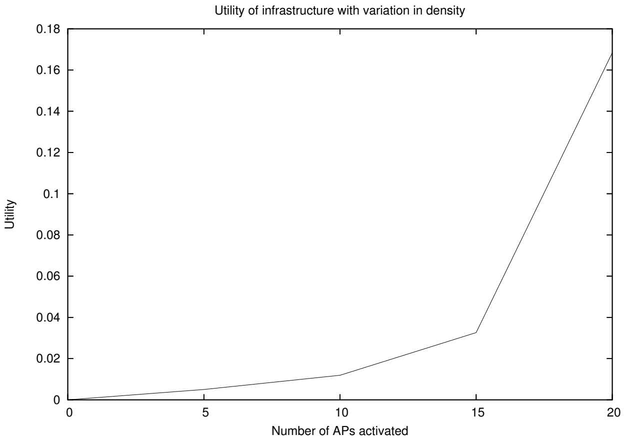
The general shape of Figure 5(a) is different from that of Figure 5(b), and Figure 5(c). We believe that this is due to the smaller number of access points in *Infocom06* dataset (20) as compared to the *Reality* (253) and *Kaist* (2500) datasets. We expect the utility of the infrastructure to further increase with more APs deployed, and eventually go through another phase transition and flat out as the other two cases.

B. Data Push

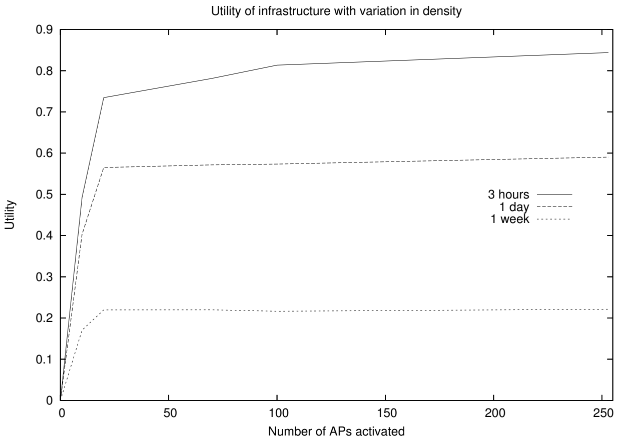
In our simulations, we assume Poisson arrival rate for the messages for each user, and set the parameters such that an average of 15 messages are generated for each person during the duration of the experiment.

When users arrive at an AP, they can pick up their own messages and (if opportunistic communication is being used) messages for other nodes from the AP. We evaluate the three forwarding schemes *APs and Opportunistic Flooding*, *APs and Opportunistic MCP*, and *Only APs* in this section. As the data push service requires messages to be originated from the infrastructure, there is always at least one access point available and there cannot be a setting where only opportunistic communication is used.

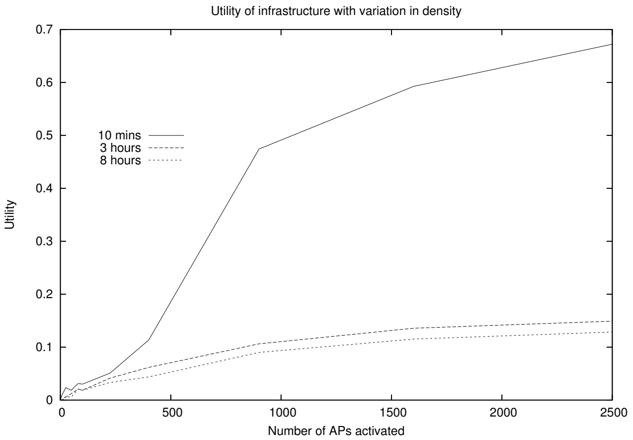
Figure 6(a) shows the delivery ratio of the data push service with the three forwarding schemes for the *Infocom06* dataset. We can see clearly from the graph that we observe a similar trend as the asynchronous messaging application, but with an



(a) *Infocom06*

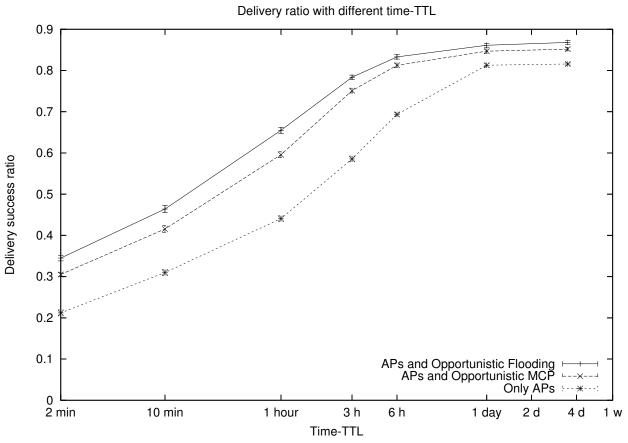


(b) *Reality*

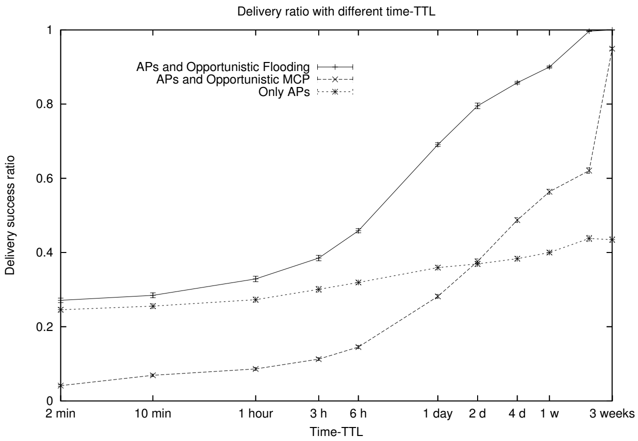


(c) *Kaist*

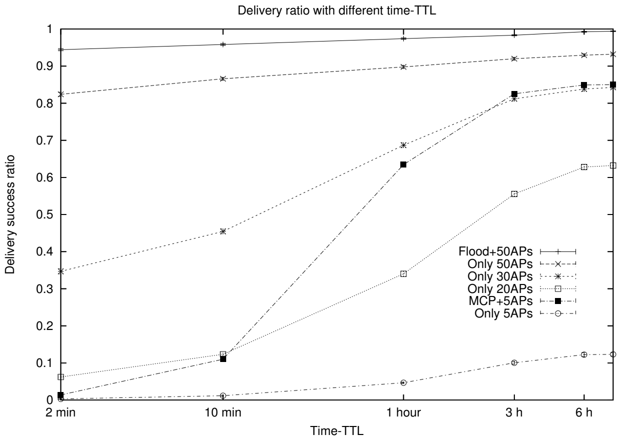
Fig. 5. Utility of infrastructure for asynchronous messaging with coexistence of opportunistic communication



(a) *Infocom06*



(b) *Reality*



(c) *Kaist*

Fig. 6. Delivery ratio for data push with time TTL

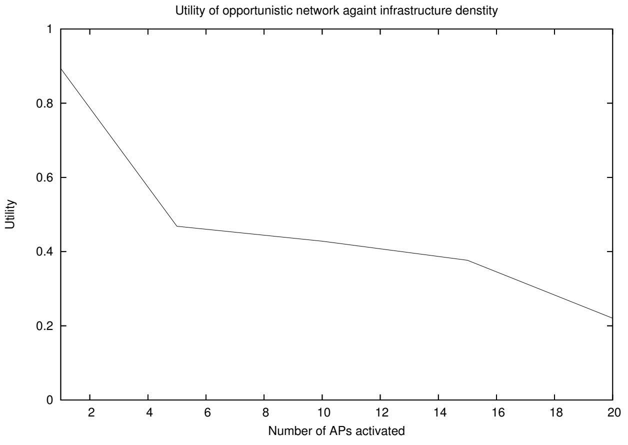
average of around 20% higher throughput. Figure 6(b) shows the delivery for the *Reality* case. The delivery ratio are also in general higher than the asynchronous messaging counter parts. Again, most of the short-delay deliveries are owing to the deliveries by infrastructure. And in this case, flooding performs much better than MCP for most cases. We choose a slightly different set of schemes to show results for in the *Kaist* case. In Figure 6(c), we try to show the change of deliveries with the number of APs and also the combined performance of opportunistic networking and infrastructured networks. In general the deliveries are also higher than the asynchronous messaging application, but the delivery ratio is still low with small number of APs. Interestingly, the combination of using opportunistic networking with only 25 ($5 \times 5 = 25$) APs outperforms (especially when the time TTL equals to 1hour and larger) the deployment of a total of 400 APs uniformly in the environment. The other properties are quite similar to the asynchronous messaging case.

A similar trend as asynchronous messaging for the utility of the opportunistic communication $U(O)$ is observed in Figure 7(a), Figure 7(b),and Figure 7(c). In general, the utilities of the opportunistic network in this case are lower than that in the asynchronous messaging scenario because the data push application need the infrastructure to provide the data source. The phase transition points also slightly shift to the left, which means that the utilities of the opportunistic networks saturate with less number of APs.

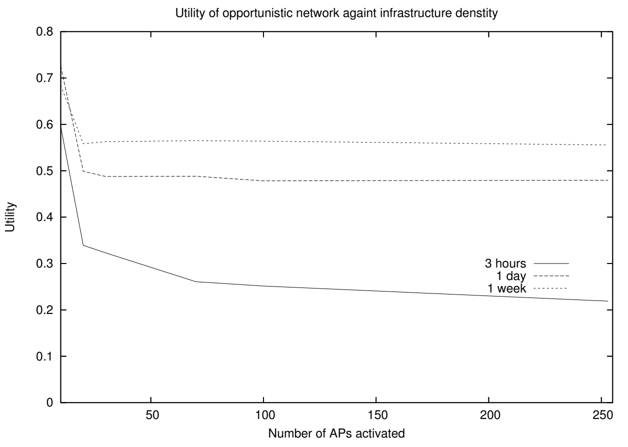
V. RELATED WORK

Lately the research area of Delay Tolerant Networks (DTN) and intermittently connected networks, has grown tremendously. Lots of work has been published on various subfields of this research area. Most of that work has however been geared towards presenting new communication architectures or protocols, and less on more fundamental analysis of underlying issues, though some such research exist. Previous work has also introduced a number of different kinds of infrastructure to support the protocol or communication system presented in that work. This infrastructure has however mostly been specialized for a particular use case. Our work is significantly different from all of the previous work. Instead of proposing the use of one particular type of infrastructure to solve a specific task, we analyze what the possible impact of introducing infrastructure into the network can be, as well as how opportunistic communication can be used to improve legacy infrastructure networks. This section gives a brief overview of some related work in the areas of mobility pattern analysis and infrastructure for sparse mobile networks.

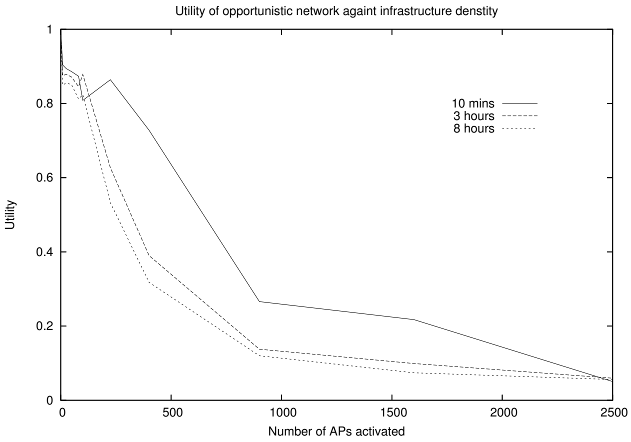
The properties of mobile users accessing an infrastructure network have been studied in the past. It has been shown that the time spent, or the amount of data transferred, during a session follow a heavy tail distribution [1]; it has been recently noted that the same result applies to the time elapsed between two sessions of a user [13]. The popularity of different locations, and the patterns of association of a typical user were also studied (see for example [4]), again exhibiting distribution



(a) *Infocom06*



(b) *Reality*



(c) *Kaist*

Fig. 7. Utility of opportunistic communication for data push with coexistence of infrastructure

with heavy tails. It has been recently proposed to include these features in several mobility models [13],[28]), or to take advantage of them for efficient location predictions, with different application in mind (see [9], [23], and references therein).

Chaintreau et al.[3] observed a power-law property in the distribution of inter-contact times in a number of experimental data sets. The authors mathematically proves that this property will cause certain simple stateless forwarding algorithms to have unbounded worst case delays. This is a very important result as it highlights the heavy-tailed nature of the inter-contact times generated by human mobility. This is very different from for example the exponential inter-contact times generate by the popular random way-point model used in many simulation studies. Extending that work, Lindgren et al.[18] made an initial study of the impact of introducing infrastructure into an intermittently connected network. That work demonstrate that, even when infrastructure is deployed to support opportunistic transfer, heavy tails are found in the time spent waiting for the next useful connection opportunity. This extends the observations made both in [3] and the studies of campus WLANs, cited above. However, it was still possible to directly measure the impact of these distributions on the delay experienced on average, and prove that under some conditions, it can still be kept reasonably short.

Our previous work [12] started a preliminary analysis of the utility of the opportunistic network with the coexistence of infrastructure and vice versa. The work was limited to a single data set, *Infocom06*, and hence less strong general conclusions could be drawn from it. In this paper, we use data from three experiments in completely different environments with both nonuniform and uniform deployments of infrastructures. The scale of the infrastructure is up to several thousand access points, and much large than in the previous paper. These provide us with more convincing evaluation results.

Banerjee *et al.* study the impact of infrastructures (relays, base stations, and meshes) density on the improvement of delivery delay of hybrid mobile network by both emulation using vehicle network traces and analytical modeling [2]. Empirically, they show that increasing the packet generation load from 5 packets per hour per destination to 120 packets per hour per destination (a factor of 24) leads to a 29%, 12%, and 28% increase in average packet delivery delay for base stations, meshes, and relays, respectively. Analytically, they also show that for N mobile nodes, the network needs $\omega(N)$ relays and $\omega(\sqrt{N})$ base stations before the stationary nodes substantially affect the average packet delivery delay for epidemic routing. Our work is different as our experimental networks are much larger in sizes, consist of human networks instead of vehicle networks, and are three different environments for generality instead of one. We also study the impact of the opportunistic network on improving the infrastructure instead of only the single side in [2] and propose the utility metric for evaluation the utility of each system. In addition, we also further verify the $\omega(\sqrt{N})$.

Liu *et al.* [22] study the capacity of wireless ad hoc networks with infrastructure support of an overlay of wired base stations using a analytical approach. They find that a two-dimensional square (or disk) network requires a large number of base stations $b = \Omega(\sqrt{N})$ before the capacity can increase effectively, where b is the number of base stations and N is the number of mobile nodes. In previous work by Liu *et al.* [21], they also obtain analytical expressions for the throughput capacity. For a hybrid network of n nodes and m base stations, the results show that if m grows asymptotically slower than $\sqrt{n}$, the benefit of adding base stations on capacity is insignificant. However, if m grows faster than $\sqrt{n}$, the throughput capacity increases linearly with the number of base stations, providing an effective improvement over a pure ad hoc network. In order to achieve non-negligible capacity gain, the investment in the wired infrastructure should be high enough. Dousse *et al.* [6] study a large-scale wireless network, but with a low density of nodes per unit area. They find that the introduction of a sparse network of base stations does significantly help in increasing the connectivity, but only when the node density is much larger in one dimension than in the other. These three works match our observation in this paper in the way that a large number of infrastructure nodes are needed to bring significant improvement on the delivery. Different from the above works, our paper focus more on the opportunities enhanced by increasing the number of contact instead of considering channel properties and congestion. This makes sense in a sparse low-contention network like *DTN*.

Several efficient forwarding algorithms for DTNs have been proposed. A majority of the algorithms are based on epidemic routing protocols [29], where messages are simply flooded when a node encounters another node. The optimisation of epidemic routing by reducing the number of copies of the message has been explored. For example, in [27], spray and wait routing assigns a limited number of copies. Many approaches calculate the probability of delivery to the destination node, where the metrics are derived from the history of node contacts, spatial information and so forth. The pattern-based Mobyspace Routing by Leguay *et al* [17], location-based routing by Lebrun *et al* [16], context-based forwarding by Musolesi *et al* [24] and PRoPHET Routing by Lindgren *et al* [20] fall into this category. PRoPHET uses past encounters to predict the probability of future encounters. The transitive nature of encounters is exploited, where indirectly encountering the destination node is evaluated. Message Ferry by Zhao *et al.* [30] takes a different approach by controlling the movement of each node.

VI. CONCLUSIONS AND FUTURE WORK

In this paper, we study the utility of opportunistic communication systems with the co-existence of network infrastructure. We observed phase transitions as an initial deployment of access points had a significant impact on the network, but after a certain point, the benefits of additional infrastructure deployments are minor and the utility of the opportunistic communication system remains stable. We also observe that with the participation of all mobile users for relaying messages for other users, the utility of a sparse infrastructure for message delivery is relatively low. We can conclude that opportunistic communication has significant possibilities of improving the system capacity and delay even with the coexistence of non-negligible amounts of infrastructure in the environment. If we know how to make use of it, we can save a lot of money from building a lot of expensive infrastructures.

In this paper, we have only focused on analysing the throughput and delay of the messages as indicators of the impact of the changes in the system. When opportunistic communication is used, there will be additional system costs induced such as additional copies of messages in the system and increased power consumption of the mobile devices. On the other hand, a system with a higher infrastructure density will have higher monetary costs as more access points must be deployed. These issues and the related tradeoffs are however not within the scope of this paper, but are interesting topics for future work. We would also like to do more extensive analysis, not only looking at other metrics, but also analysing data from other connectivity traces to make sure the results are generally applicable.

In the future, we also want to have deeper comparison of our observations with the theoretical works above([22] [21] [6]). From this paper, we can also see that uniform deployment is not efficient, and we want to look for more cost-effective way of infrastructure deployment strategies.

VII. ACKNOWLEDGEMENTS

This research is funded in part by the EC IST SOCIALNETS - Grant agreement number 217141. We would like also to acknowledge the members of the Haggle project and the NCSU for the connectivity traces.

REFERENCES

[1] BALAZINSKA, M., AND CASTRO, P. Characterizing mobility and network usage in a corporate wireless local-area network. In *Proceedings of the First International Conference on Mobile Systems, Applications, and Services (MobiSys 2003)* (2003).

[2] BANERJEE, N., CORNER, M. D., TOWSLEY, D., AND LEVINE, B. N. Relays, Base Stations, and Meshes: Enhancing Mobile Networks with Infrastructure. In *Proceedings of ACM MobiCom* (San Francisco, CA, USA, Sept. 2008).

[3] CHAINTREAU, A., HUI, P., SCOTT, J., GASS, R., CROWCROFT, J., AND DIOT, C. Pocket switched networks: Real-world mobility and its consequences for opportunistic forwarding. Tech. Rep. UCAM-CL-TR-617c, University of Cambridge, Computer Lab, February 2005.

[4] CHINCHILLA, F., LINDSEY, M., AND PAPADOPOULI, M. Analysis of wireless information locality and association patterns in a campus. In *Proceedings of IEEE International Conference on Computer Communications (INFOCOM'04)* (2004).

[5] CRUCITTI, P., LATORA, V., AND PORTA, S. Centrality measures in spatial networks of urban streets. *Physical Review E 73* (2006), 036125.

[6] DOUSSE, O., THIRAN, P., AND HASLER, M. Connectivity in ad-hoc and hybrid networks. In *in Proc. IEEE Infocom* (2002), pp. 1079–1088.

[7] EAGLE, N., AND PENTLAND, A. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing V10*, 4 (May 2006), 255–268.

[8] FALL, K. A delay-tolerant network architecture for challenged internets. In *Proc. SIGCOMM* (2003).

[9] HENDERSON, T., KOTZ, D., AND ABYZOV, I. The changing usage of a mature campus-wide wireless network. In *Proceedings of the Tenth Annual International Conference on Mobile Computing and Networking (MobiCom 2004)* (2004).

[10] HUI, P., CHAINTREAU, A., ET AL. Pocket switched networks and human mobility in conference environments. In *Proc. WDTN* (2005).

[11] HUI, P., CROWCROFT, J., AND YONEKI, E. Bubble rap: Social-based forwarding in delay tolerant networks. In *MobiHoc '03: Proceedings of the 9th ACM international symposium on Mobile ad hoc networking & computing* (May 2008).

[12] HUI, P., AND LINDGREN, A. Phase transition of opportunistic communications. In *Proceedings of the ACM Workshop on Challenged networks (CHANTS)* (Sept. 2008).

[13] JAIN, R., SHIVAPRASAD, A., LELESCU, D., AND HE, X. Towards a model of user mobility and registration patterns. *Mobile Computing and Communications Review 8*, 4 (2004), 59–62.

[14] JANSON, S., LUCZAK, T., AND RUCINSKI, A. *The Phase Transition.* New York: Wiley, 2000, ch. 5 in Random Graphs, pp. 103–138.

[15] JUPITERRESEARCH. Municipal wireless: Partner to spread risks and costs while maximizing benefit opportunities. Tech. rep., Jupitermedia Corporation, 2005.

[16] LEBRUN, J., CHUAH, C.-N., ET AL. Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks. *IEEE VTC 4* (2005), 2289–2293.

[17] LEGUAY, J., FRIEDMAN, T., ET AL. Evaluating mobility pattern space routing for DTNs. In *Proc. INFOCOM* (2006).

[18] LINDGREN, A., DIOT, C., AND SCOTT, J. W. Impact of communication infrastructure on forwarding in pocket switched networks. In *Proceedings of the ACM SIGCOMM workshop on Challenged Networks (CHANTS 2006)* (September 2006).

[19] LINDGREN, A., AND DORIA, A. Experiences from a deploying a real-life DTN system. In *Proceedings of the 4th Annual IEEE Consumer Communications and Networking Conference (CCNC 2007)* (January 2007).

[20] LINDGREN, A., DORIA, A., ET AL. Probabilistic routing in intermittently connected networks. In *Proc. SAPIR* (2004).

[21] LIU, B., LIU, Z., AND TOWSLEY, D. On the capacity of hybrid wireless networks. In *in Proc. IEEE Infocom* (2003), pp. 1543– 1552.

[22] LIU, B., THIRAN, P., AND TOWSLEY, D. Capacity of a wireless ad hoc network with infrastructure. In *MobiHoc '07: Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing* (New York, NY, USA, 2007), ACM, pp. 239–246.

[23] MCNETT, M., AND VOELKER, G. M. Access and mobility of wireless pda users. *Mobile Computing Communications Review 9*, 2 (April 2005), 40–55.

[24] MUSOLESI, M., HAILES, S., ET AL. Adaptive routing for intermittently connected mobile ad hoc networks. In *Proc. WOWMOM* (2005).

[25] RHEE, I., SHIN, M., HONG, S., LEE, K., AND CHONG, S. On the levy-walk nature of human mobility. In *Proc. of IEEE INFOCOM* (Phoenix, USA, April 2008).

[26] SASTRY, N., SOLLINS, K., AND CROWCROFT, J. Architecting citywide ubiquitous wi-fi access. In *Proceedings of The Sixth Workshop on Hot Topics in Networks (HotNets-VI)* (2007).

[27] SPYROPOULOS, T., PSOUNIS, K., ET AL. Spray and wait: An efficient routing scheme for intermittently connected mobile networks. In *Proc. WDTN* (2005).

[28] TUDUCE, C., AND GROSS, T. A mobility model based on wlan traces and its validation. In *Proceedings of IEEE International Conference on Computer Communications (INFOCOM'05)* (2005).

[29] VAHDAT, A., AND BECKER, D. Epidemic routing for partially connected ad hoc networks. Tech. Rep. CS-200006, Duke University, April 2000.

[30] ZHAO, W., AMMAR, M., AND ZEGURA, E. A message ferrying approach for data delivery in sparse mobile ad hoc networks. In *Proceedings of the MobiCom 2004* (2004).