

Opportunistic Content Distribution in an Urban Setting

J r mie Leguay*, Anders Lindgren+, James Scott , Timur Friedman*, Jon Crowcroft#

*Universit  Pierre et Marie Curie, LiP6-CNRS

+Lule  University of Technology

 Intel Research Cambridge

#University of Cambridge

ABSTRACT

This paper investigates the feasibility of a city-wide content distribution architecture composed of short range wireless access points. We look at how a target group of intermittently and partially connected mobile nodes can improve the diffusion of information within the group by leveraging fixed and mobile nodes that are exterior to the group. The fixed nodes are data sources, and the external mobile nodes are data relays, and we examine the trade off between the use of each in order to obtain high satisfaction within the target group, which consists of data sinks. We conducted an experiment in Cambridge, UK, to gather mobility traces that we used for the study of this content distribution architecture. In this scenario, the simple fact that members of the target group collaborate leads to a delivery ratio of 90%. In addition, the use of external mobile nodes to relay the information slightly increases the delivery ratio while significantly decreasing the delay.

Categories and Subject Descriptors

C.2 [Computer-Communication Networks]: Network Architecture and Design

General Terms

Measurement, Experimentation, Performance, Algorithms

Keywords

Content distribution, Mobility data, Delay Tolerant Networking

1. INTRODUCTION

This increased penetration of wireless-capable handheld devices has led to the development of new communication techniques. Such communication techniques include *opportunistic networking*, which makes use of the capability of the devices to communicate locally among their neighbors

to create communication possibilities with users and devices in other places, even when if there never exists a fully connected path between the two end-points. These networks are a type of delay tolerant network (DTN) [4] and fall also under the Pocket Switched Networking (PSN) paradigm [2]. In this context, this paper investigates the feasibility of a city-wide content distribution architecture for electronic newspapers or local information. We look at how a target group of intermittently and partially connected mobile nodes can improve the diffusion of information within the group by leveraging various mixtures of fixed and mobile nodes that are exterior to the group. The fixed nodes are data sources, and the external mobile nodes are data relays, and we examine the trade off between the use of each in order to obtain high satisfaction within the target group, which consists of data sinks.

To evaluate the different content distribution schemes we propose, we conducted an experiment in the city of Cambridge, UK, in which 20 stationary devices equipped with a Bluetooth contact logger were deployed at popular places. We then ran simulations in which we imagined that these devices were access points distributing electronic content. In addition to this, we deployed 40 similar contact loggers on a group of students from Cambridge University. Because we used Bluetooth technology, we gathered interactions not only between the contact loggers, but also with a large number of other Bluetooth enabled devices such as mobile phones or PDAs. In our simulations, students were the target group, making the assumption that they were all interested in the content distributed by the access points, and Bluetooth devices external to the experiment could potentially be data relays. We are making the data collected in this experiment available to the research community [3]. We therefore devote a part of this paper to a description of the salient characteristics of the dataset.

This paper has two main contributions. First, it presents an original data set using fixed iMotes. Second, using these data, it evaluates performance of a city-wide content distributing architecture. This paper validates the use of opportunistic networking in the particular environment we studied. It shows that despite the fact that students did not on average meet a large number of access points each day, we can achieve good performance in delivery ratio, delay and resource utilization with a content distribution scheme that allows students to collaborate. We also demonstrate that the use of Bluetooth devices external to the experiment to relay the content can make an incremental but important increase in performance in both an increased delivery ratio

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

SIGCOMM'06 Workshops September 11-15, 2006, Pisa, Italy.
Copyright 2006 ACM 1-59593-417-0/06/0009 ...\$5.00.

and a decreased delay. Finally, we investigate the robustness of the content distribution infrastructure and we show that decreasing the number of members of the target group or the number of access points increase interest of using exterior nodes as relays.

The rest of this paper is structured as follows. Sect. 2 describes the experiment setup. Sect. 3 presents the analysis of the mobility traces that were collected. Sect. 4 details the content distribution schemes proposed and evaluates them. Sect. 5 provides an overview of related work concerning mobility data acquisition. Sect. 6 concludes the paper.

2. EXPERIMENT SETUP

In the experiment we performed, we were interested in tracking contacts between different mobile users, and also contacts between mobile users and various fixed locations. Previous experiments have measured contacts between mobile users in corporate and conference settings [2] by requesting users to carry small Intel iMote¹ devices that can log contacts with other Bluetooth enabled devices. We chose to use the same technology to gather contacts. Mobile users in our experiment mainly consisted of students from Cambridge University who were asked to carry these iMotes with them at all times for the duration of the experiment. In addition to this, we deployed a number of stationary nodes in various locations that we expected many people to visit such as grocery stores, pubs, market places, and shopping centers in and around the city of Cambridge, UK. A stationary iMote was also placed at the reception of the Computer Lab, in which most of the experiment participants are students. Figure 1 shows the positions of the stationary nodes. The road that rings the center of Cambridge, an area of 3 km², is clearly visible on the map.



Figure 1: Locations of fixed iMotes.

To discover other nearby users and to be able to log contacts between nodes, the iMotes use the Bluetooth inquiry mechanism that allows them to get knowledge of all other Bluetooth enabled devices within radio transmission range. As conducting the inquiries requires transmitting and receiving over the radio interface, this consumes power and a trade-off that had to be considered when setting up the experiment was how to set δ , the interval between inquiries. Indeed, having a δ too low would have led to a shortened lifetime of the iMotes due to the high power consumption from frequent use of the radio. On the other hand, setting δ to a too high value means running the risk of missing more

¹The iMotes are small sensor platforms with an ARM7 processor and some on board storage and Bluetooth capability.

potential contacts. Note that when an iMote is not inquiring, it answers to other iMotes' enquiries.

To determine the inquiry interval to use, we studied power consumption on the iMotes while idle and while performing inquiries. Using these measurements in conjunction with experience on the life-time of iMotes in previous experiments, we chose inquiry intervals that we hoped would allow the devices to have a life-time of 2 weeks. Furthermore, there is a small risk that the Bluetooth inquiry may occasionally miss a contact even though it is present. Therefore, we made the decision that if a contact is seen at a given inquiry I_i , but not at the subsequent at inquiry I_{i+1} , we will still assume that the recorded contact was never broken if we observe it again at the following inquiry I_{i+2} . This assumption was also made in previous contact logging experiments using iMotes.

iMotes carried by students had to be packaged within a small form factor to increase the probability that the users would actually always carry the device and not leave it behind. On the other hand, we had larger freedom when it came to the stationary devices. Thus, for some of the fixed iMotes, we added extra battery power to be able to reduce the inquiry interval so that we would detect more of the possible contacts. Furthermore, on a few of the fixed iMotes, we were also able to attach external antennas with greater wireless range. This increases the coverage area in which they can detect mobile devices in large public places.

- *MSR-10*: Mobile Short Range iMotes with an interval of 10 minutes between inquiries. These iMotes were given to a group of 40 students, mostly in the 3rd year at the Cambridge University Computer Lab. The devices were packaged in small boxes (dental floss boxes) to be easy to carry around in a pocket, and used a CR-2 battery (950 mAh) for power.
- *FSR-10*: Fixed Short Range iMotes with an interval of 10 minutes between inquiries. We deployed 15 of these iMotes in fixed locations such as pubs, shops or colleges' porter lodge. We used exactly the same packaging and batteries as the MSR-10.
- *FSR-6*: Fixed Short Range iMotes with an inquiry interval of 6 minutes. These iMotes were equipped with a more powerful rechargeable battery providing 2200 mAh so that we were able to reduce the inquiry interval to 6 minutes. We deployed 2 of these.
- *FLR-2*: Fixed Long Range iMotes with an interval of 2 minutes between inquiries. To increase the area in which these iMotes can discover other devices, four devices were equipped with an external antenna, which provided a communication range that was approximately twice that of the short range iMotes. Further, these iMotes were also equipped with 3 more powerful rechargeable batteries providing 2200 mAh so that we could reduced the inquiry interval to 2 minutes. Their antenna and packaging can be seen in Fig. 2.

To prevent the results from being biased by the fact that the mobile devices are co-located as they were being deployed to their carriers, we have removed the data collected during first 3 hours of the experiment from the analysis. After the mobile devices had been given to the experiment participants, we proceeded to the city centre to deploy the stationary iMotes at their respective locations. The experiment



Figure 2: Long range iMote with rechargeable batteries.

started on Friday, October 28th 2005, 9:55:32 (GMT) and stopped on Wednesday, December 21th 2005, 13:00 (GMT).

3. DATA ANALYSIS

Due to various hardware problems and the loss of some of the deployed iMotes, we were able to gather measurement data from 36 mobile participants and 18 fixed locations, as Table 1 shows. This table presents statistics about the experiment. It shows that the average lifetimes for all types of iMotes is higher than 10 days and that these results present a low variability by type except for FLR-2. Indeed, while 2 of the FLR-2 could remain active for the full 23 days of the deployment, 2 of them that were deployed in very popular places suffered from a buffer overflow after 5 and 9 days respectively, having recorded on average 3,670 contacts.

In our analysis, we consider two categories of contacts: *internal contacts*, which are contacts that occurred between two iMotes of any type (fixed or mobile), and *external contacts*, which are contacts that occurred between an iMote and another Bluetooth capable device (e.g., PDA or mobile phone). Table 1 shows the number of contacts acquired by all the types of iMotes for all the categories of contacts. The table also lists the number of *unique* contacts that has been seen in the different categories. Unique contacts are the number of different node pairs that ever have a contact over the course of the experiment duration. We can first see that, as expected, the MSR-10 iMotes had a large number of contacts with each other and that they also had a significant number of contacts with external devices, 10,469 in total. The second immediate observation is that fixed iMotes had a very large number of contacts (20,240 in total) with external devices, while they did not meet the participants of the experiment very much, with only 231 contacts in total. Despite the small number of FLR-2 iMotes that were deployed, their placement at very popular locations allowed them to capture a large number of external contacts.

3.1 Inter-students contacts

Here we analyse the interactions we observed between participants carrying iMotes. We first note that they had a large number of contacts together, as Fig. 3 shows. Fig. 3(a) and Fig. 3(b) present, respectively, for each mobile iMote, the number of total contacts and unique contacts. On average, students had 461.9 internal contacts with a standard deviation of 196.2 and 30.0 unique internal contacts over the 35 possible with a standard deviation of 4.0.

Fig. 4 shows the number of contacts per day between mobile iMotes. Most of the contacts occurred during week days,

	MSR-10	FSR-10	FSR-6	FLR-2
Nb motes	36	12	2	4
Lifetime (days)	10.7 \pm 0.8	11.0 \pm 0.6	14.5 \pm 0.5	15.7 \pm 8.3
Contacts	19014	8270	1082	11119
Int. co.	8545	38	91	102
Ext. co.	10469	8232	991	11017
Contacts (u)	5681	6189	815	6789
Int. co. (u)	644	25	35	43
Ext. co. (u)	5037	6164	780	6746

Table 1: Global statistics. (u) means unique contacts.

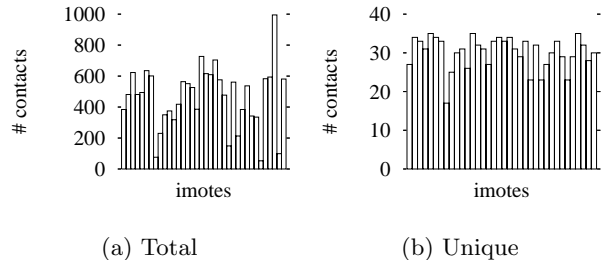


Figure 3: Contacts between mobile iMotes.

and less contacts have been recorded during Saturdays and Sundays (i.e., days 2, 3 and 9, 10). In a group of students, in which most of them belong to the same program, this observation is natural.

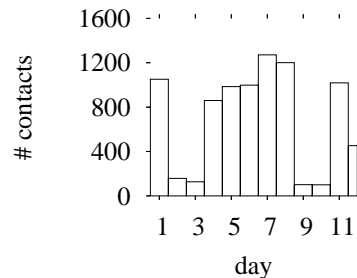


Figure 4: Number of contacts per day between mobile iMotes.

In Fig. 5 we can see the distribution of the inter-contact time between students. The inter-contact time is the time between two contacts for a given node pair, and its distribution has previously been shown to exhibit a power-law behavior in a large number of experiments[2]. We see similar tendencies to power-law behavior as in previous experiments here, but we can see that a large part (over 90%) of the inter-contact times are shorter than one day. This means that after a node pair have met, there is a 90% chance they will meet again within one day. The exponent of the power law is equal to 0.46.

3.2 Contacts with fixed iMotes

As explained previously, one of the goals of the experiment we present in this paper was to explore not only the interaction between the participants wearing iMotes but also

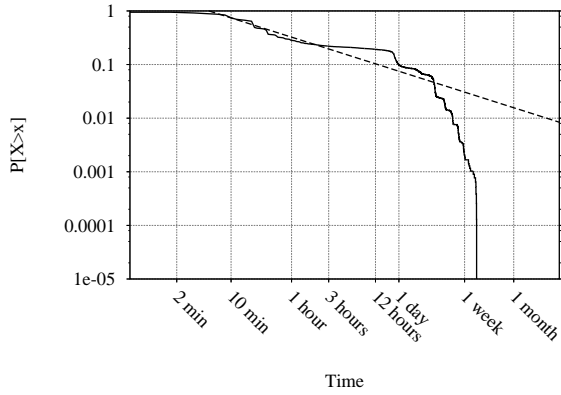


Figure 5: Inter-contact time distribution

to capture their mobility from fixed locations distributed at popular places in the city. However, the results we obtain do not meet our expectations as shown by the plots in Fig. 6. Fig. 6(a) and Fig. 6(b) present, for each of the fixed iMotes, the total number of contacts with mobile iMotes and the number of unique mobile iMotes observed, respectively. They show in detail that very few contacts occurred between iMotes carried by students and most of the fixed ones. The only two fixed iMotes having significantly more contacts with students were those at the reception at the Computer Lab (where the students attend class activities) and at a popular grocery store.

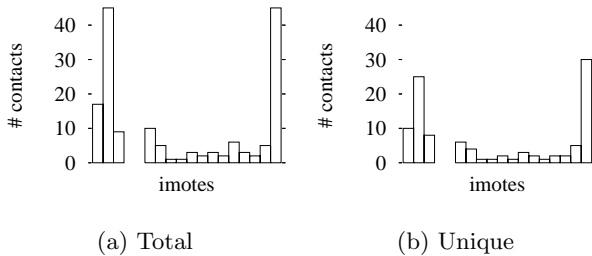


Figure 6: Contacts between mobile and fixed iMotes.

There are a number of factors that can explain this result. First, it might be possible that the fixed iMotes were deployed at inappropriate locations according to the population sample. Before the deployment, an attempt was made to survey students about popular locations to visit, and this in conjunction with reasoning on where people are likely to go (which is possible in a city of Cambridge’s size), the locations were chosen. Apparently, students did not experience a large number of contacts with locations where we had deployed the iMotes. As we will see in the next section, the fixed iMotes did however log many external contacts, verifying that the locations in which they were deployed were indeed frequently visited by people with Bluetooth enable devices – just not experiment participants. This kind of deployment might work better in corporate environments in which people are confined all the day or in experiments

with more participants. Second, we might have missed logging many contacts, especially in transit areas. This issue is discussed later in Sect. 3.4.

3.3 External contacts

In addition to measuring contacts between iMotes, all contacts between the iMotes and other Bluetooth enabled external devices were also logged. While this was not the main objective of the experiment, this data ended up constituting the largest part of our data set. Indeed, we observed 10,469 contacts (3,586 unique) between mobile iMotes and external devices, and 20,240 contacts (9,211 unique) between fixed iMotes and external devices. Here we investigate these contacts with external devices by first quantifying them and then trying to identify the nature of these devices.

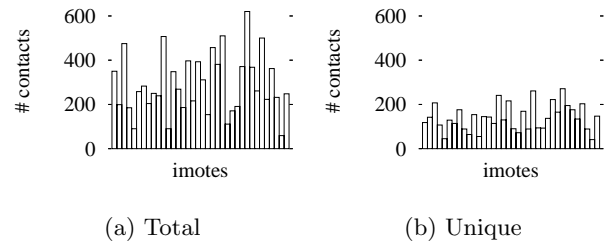


Figure 7: Contacts between external devices and mobile iMotes.

Fig. 7 shows the contacts each mobile iMote had with external devices. Fig. 7(a) and Fig. 7(b) show the total number of external contacts and the number of unique external contacts respectively. Mobile iMotes acquired on average 290.8 external contacts and 139.9 unique external contacts with respectively a standard deviation of 132.6 and 139.9.

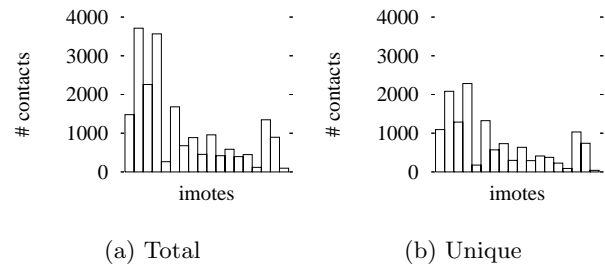


Figure 8: Contacts between external devices and fixed iMotes.

Fig. 8 is similar to Fig. 7 but for the fixed iMotes. In that case, fixed iMotes acquired on average 1124.7 external contacts with a standard deviation of 1049.7 and 760.5 unique external contacts with a standard deviation of 632.3. The fact that the number of contacts is higher for the 4 first days is due to an iMote that ran rapidly out of memory, being placed in a very popular location. Note that we did not consider this iMote in simulations presented next section.

To continue our investigation on external devices, we used

the database of Organizationally Unique Identifiers (OUI)² maintained by the IEEE to map MAC address prefixes in the data set to their manufacturers. We were able to resolve 97% of the prefixes. Fig. 9 presents the frequency of the occurrence of the most common manufacturers. From looking at the manufacturers, we can see that most of external devices are likely mobile phones or other portable devices (Murata is a Bluetooth chip manufacturer whose products are integrated in a wide range of devices such as mobile phones, Personal Digital Assistants (PDAs), laptops, etc.), as opposed to devices such as printers or wireless keyboards and mice. This observation is of great interest because we can reasonably consider external devices in the data set as mobile entities that are carried in pockets of regular people moving around the city, exactly as mobile iMotes.

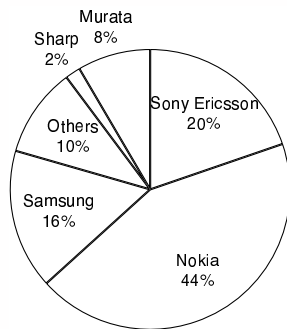


Figure 9: Manufacturers

3.4 Discussion

When using devices like the iMotes to gather data about contact patterns, measurements may not be able to exhaustively capture contacts because of two main factors. First, as these experiments require the active involvement of participants, there is always the risk that the participants do not completely fulfill their commitments, for example by not always carrying the measurement devices everywhere they go. Indeed, as we conducted a survey on our population after the experiment, we were able to determine that some of them had occasionally forgotten to bring the iMote when going somewhere, or might have left it in a bag instead of keeping it on their person. There were also occasions of students leaving the city over the weekend (which is less of a problem as that reflects a real user behavior, and will still be able to gather external contacts, but most likely no internal contacts). Secondly, as contacts are only discovered using the periodic Bluetooth inquiries, iMote experiments suffer from a sampling effect that means that contacts that are shorter than δ minutes may be missed. This is a trade-off between minimizing the risk of missing short contacts and the life-time of the iMotes. It would be possible to extend the life-time further while keeping a short inquiry interval by adding more powerful batteries, but that would result in a more bulky form factor.

4. CITY-WIDE CONTENT DISTRIBUTION

²<http://standards.ieee.org/regauth/oui/>

4.1 Scenario

We propose here an evaluation of a city-wide content distribution architecture. As said previously, this architecture is composed of wireless short range access points disseminated down town at popular places. Content (e.g., newspapers or local information) is opportunistically distributed to nodes that pass close to these access points. We consider fixed iMotes we deployed in Cambridge to be the content distribution access points and mobile iMotes given to students being the target group. Our aim here is to propose and evaluate schemes that distribute content to a population of users interested to which our target group is assumed to belong (i.e., students from Cambridge University wearing mobile iMotes). Members of the target group were not aware of the positions of access points.

In this scenario, access points generate a new copy of a given newspaper at 7 am every day. Once acquired, copies are kept by nodes till 7 am the next day. We have replayed in this evaluation 5 days of data gathered in Cambridge from Monday to Friday. We removed the fixed iMote at the Computer Lab because it was located at the place the community of student met most of the time (we did not want to reduce the study to a trivial exercise).

4.2 Distribution schemes

Within the scenario previously described, we evaluated the following distribution schemes:

- *Selfish*: nodes get the content directly from the access points and never pass it on to other nodes. The access points distribute an unlimited number of copies.
- *Collectivist*: nodes can get content directly from access points and are able to share it within the communities they belong to. Note that nodes in our target group are assumed to belong to the same community.
- *Extended collaboration*: in addition to *Collectivist* strategy, external mobile devices can be used to relay the content. The details of the schemes we propose are presented later in this section. External mobile devices may be of several kinds. In the case of a collaborative scheme may involve strangers, mechanisms to incite nodes to relay the content need to be provided but are not the focus of this paper. Also note that the experiment we conducted did not provided us contacts between external devices.
- *Top students*: use only the N students that had the highest number of contacts to be able to pass copies to the others.
- *Strangers only*: students can not relay the content, only external devices are used as relays.

In order to define heuristics to select the external devices that would be involved in schemes using extended collaborations, we obtained statistics on potential *mobile bridges*, defined as nodes having seen at least a fixed and a mobile iMote during the experiment. These nodes represent 12.5% (1,430 over 11,367) of external devices, they are potentially interesting to act as relays between access points and the targeted population.

In Fig. 10, we plot one point for each of the mobile bridges, showing the number of contacts it had with fixed and mobile

iMotes. We add some small random noise in order to obtain a cloud of points. We can first observe that no mobile bridges had a large number of contacts with both fixed and mobile iMotes. A given mobile bridge seems to be close in terms of its mobility to either a fixed iMote or to a mobile one. Mobile bridges had on average 3.8 contacts with fixed iMotes with a standard deviation of 5.6 and 4.3 with mobile iMotes with a standard deviation of 19.3.

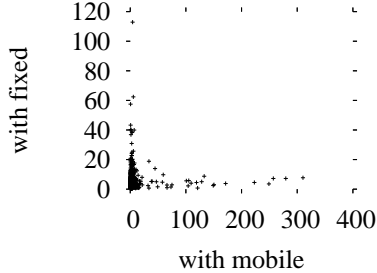


Figure 10: Number of contacts with fixed and mobile iMotes.

We define a pair here as a set of one fixed and one mobile iMotes that could be potentially covered by a mobile bridge, meaning that there is at least one mobile bridge that has seen these mobile and fixed iMotes. We found that 610 pairs between the sets of fixed and mobile iMotes exist in the data over the 680 possible pairs. To have a better idea of the strength of the coverage of pairs by mobile bridges, we plot in Fig. 11 the probability distribution that pairs are covered by more than X mobile bridges. This figure shows that some pairs are covered by a significant number of mobile bridges. 10% of the pairs are covered by more than 20 bridges.

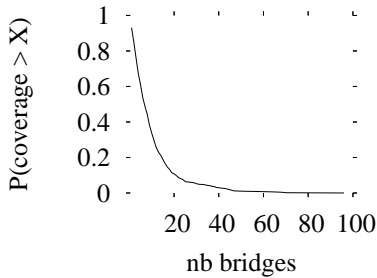


Figure 11: Coverage of pairs by mobile bridges.

All these observations concerning mobile bridges lead us to think that strong hidden connections exist between the two sets that as first seen to be largely disconnected (as seen in Sect. 3.3, the number of contacts between fixed and mobile iMotes have been observed to be low). They also motivate the definition the following variations for the extended collaboration scheme:

- *All external*: use all external devices as relays.
- *Top external*: use the N external devices that had the highest number of contacts.
- *Top coverage pairs*: use the N mobile bridges covering the highest number of pairs (fixed, mobile).

- *Top coverage contacts*: use the N mobile bridges having the highest number of contacts with both mobile and fixed iMotes. Note that for all the 3 previous schemes, an algorithm have to be defined to choose these high potential relays in a distributed fashion. We let this for further work.

4.3 Performance evaluation

To measure the performance of the different content distribution schemes, we used the following metrics:

- *Delivery ratio*: the percentage of *bundles* (from the DTN terminology, i.e., messages containing the electronic newspaper) that were delivered. In our scenario, the maximum number of bundles that could be delivered is 175 (1 bundle is expected for each of the 35 mobile nodes each of the 5 days of simulation). This metric evaluates the user satisfaction.
- *Average delay*: the average bundle delay (computed on the bundles delivered).
- *Efficiency*: the number of messages transmitted per bundle delivered. It represents a measure of the network resource usage.

	Delivery	Delay	Efficiency
Selfish	20.5	7.47	1.00
Collectivist	90.2	5.29	1.00
All external	97.1	4.10	36.4
Top external			
N=1	90.2	5.29	1.03
5	90.2	5.23	1.15
10	91.4	4.45	1.28
50	92.5	4.50	2.26
100	94.2	4.60	3.33
Top cov. pairs			
N=1	90.2	5.29	1.02
5	91.4	4.44	1.13
10	91.4	4.44	1.26
50	94.2	4.60	2.09
100	95.4	4.59	2.86
Top cov. contacts			
N=1	90.2	5.29	1.03
5	90.2	5.25	1.11
10	91.4	4.40	1.21
50	93.7	4.47	2.00
100	94.8	4.45	2.78
Top students			
N=1	20.5	7.46	1.00
5	56.5	9.52	1.00
10	66.2	7.55	1.00
35	90.2	5.29	1.00
Strangers only	66.2	8.06	40.99

Table 2: Simulation results.

Table 2 presents the simulation results. The first thing that we can observe is that the selfish strategy leads to poor results in delivery ratio (20.5%), which seems natural since we did not measure a large number of contacts between students and access points. However, we see a great improvement when the students collaborate, leading to 90.2% of delivery. Moreover, what we can see from the results regarding the extended collaboration scheme is that delivery ratio is slightly improved when increasing the number of relays selected while the delay is significantly decreased being close

to the minimum that can be achieved with our data (i.e., 4.10 hours when using all the nodes). Note that the top coverage contacts or the top coverage pairs seem to be the most efficient strategies among the ones evaluated when selecting a small number of external devices as relay. With $N=10$, we reduce the delay for the top coverage contacts by 20% while increasing the delivery ratio from 1.4% compared to the collectivist scheme. Finally, external devices seem not to be sufficient to ensure a high delivery ratio themselves. When only using strangers to relay the content we only achieve a 66.2% delivery ratio.

What the results tell here is that, collaboration inside the target group improves a lot the performance while only adding one transmission per bundle delivered compared to the selfish strategy. Furthermore, a slight gain in delay and a significant one in delivery can be achieved when using a few nodes that are *close* both to students and to access points.

Robustness with number of access points

We evaluate here the *robustness* of this content distribution infrastructure by looking at its performance if some of the most popular access points are removed (the ones with highest numbers of contacts). We removed access points in order from the most popular one to the least. Fig. 12 presents the results in delivery ratio for the following schemes: collectivist, all external and extended collaborations (top coverage pairs and top coverage contacts with $N=10$). It shows that the interest of using external devices as relays is clear when the number of access points is decreased. Extended collaborations achieve a delivery ratio in between collectivist and all external.

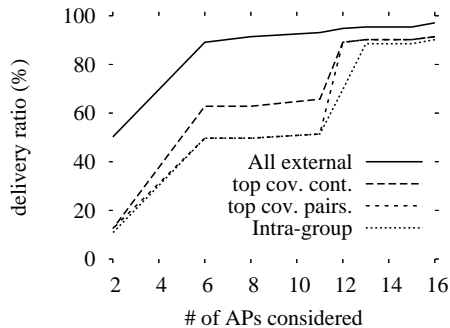


Figure 12: Delivery ratio when removing popular APs.

Robustness with number of students

Similarly, we evaluate here the *robustness* of this content distribution infrastructure by looking at its performance if most popular students are removed (the one with highest number of contacts). We removed students from the most popular to the less one. Fig. 13 presents the results in delivery ratio for the same schemes as previously. Again, we see clearly that the use of external devices increases the delivery ratio when the number of members of the community is decreased.

5. RELATED WORK

Efforts to acquire mobility data for DTN scenarios have expanded rapidly in the past couple of years. The Reality

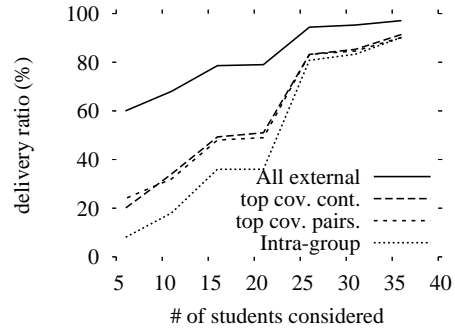


Figure 13: Delivery ratio when removing popular students.

Mining [5] experiment conducted at MIT has captured proximity, location, and activity information from 100 subjects over an academic year. Each participant had an application running on their mobile phone to record proximity with others through periodic Bluetooth scans and location using information provided by the phone on the cellular network. The UMass DieselNet project [1] also aims to study DTN routing in challenging contexts such as power outages or natural disasters. A testbed to gather interactions between 40 buses in western Massachusetts was deployed in 2005. In addition to the experiment described in the paper, other experiments with iMotes have been conducted by the Haggle [2] project, which explores networking possibilities for mobile users using peer-to-peer connectivity between them in addition to existing infrastructures. To show the similarities and differences between this and previous studies, we summarize the main parameters and measurement results from all the experiments in Table 3, extending the information provided by Chaintreau et al. [2].

The experiments *Intel* and *Cam-U* were performed in corporate and research lab settings, with the participants being researchers and graduate students. The *Infocom05* experiment was conducted at a research conference and the *Cambridge* experiment is the experiment presented in this paper. We see that this experiment spans 2-3 times as much time as previous experiments with similar number of mobile devices as the conference experiment, but with significantly more iMotes than in the first two experiments. In the other experiments, we see a high number of internal contacts while in this experiment, whereas in this experiment this number is much lower. On the other hand, the number of external devices seen and the number of contacts with them are much higher in this experiment than in previous ones. Both of these differences can be explained by the population of participants and the setting in which the experiment was deployed. In the previous experiments, participants were chosen from either people that work together on a daily basis at the same premises or attend the same conference. Thus, it is natural that they will have frequent contacts with each other. On the other hand, in this experiment, students might not have pre-existing relationships with each other and are thus less likely to have contacts outside class activities.

Other work has been to gather data that can be used, after some processing, as DTN-like data. For instance, Henderson et al. at Dartmouth College [6] have deployed one of the

	Intel	Cam-U	Infocom05	Cambridge
Duration (days)	3	5	3	10
δ (mins)	2	2	2	10
Devices	8	12	41	36
Internal contacts	1091	4229	22459	8545
Average # Contacts/pair/day	6.5	6.4	4.6	1.5
External devices	92	159	197	3586
External contacts	1173	2507	5791	10469

Table 3: Comparison with previous experiments for data from mobile iMotes.

most extensive trace collection efforts to gather information about its Wi-Fi access network. These data have been used as mobility data to characterise the mobility of users [7] or to evaluate DTN routing protocols [9]. Similar Wi-Fi based data have been used to analyse mobility such as that of ETH Zürich [13]. Furthermore, the data presented in the paper might be of great interest to evaluate forwarding algorithms defined for DTNs such as the work by Vahdat et al. [14] that uses epidemic routing, the Spray and Wait [12] protocol that distributes a number of copies to relays and then waits until the destination meets one of them, MobySpace [9] that uses a virtual space based upon nodes' mobility patterns, or the PРоPHET routing protocol [10], which bases routing on a probabilistic metric calculated using history of encounters and transitivity.

Finally, a very close scenario to that of our work was introduced in a research note by Lawrence et. al [8]. They envision to create a community content distribution network using *familiar strangers* [11], i.e. people who we meet very regularly but who we do not know. This kind of node may be present in the data we collected.

6. CONCLUSION AND FUTURE WORK

We have proposed and evaluated in this paper schemes for distribution of content in an urban environment using short range Bluetooth access points. To evaluate these schemes, we conducted a city-wide experiment using Intel iMotes, which are Bluetooth contact loggers. Stationary iMotes were deployed at popular places to act as content distribution access points while students, considered as our target group, from Cambridge University were carrying other iMotes in their pockets. We show that the simple fact that students collaborate led, in this experiment, to a delivery ratio of 90% and that the additional use of Bluetooth devices external to the experiment to relay the information slight increased the delivery ratio while significantly decreasing the delay. We have also shown that the interest of using external devices as relays increases when the size of the infrastructure and of the targeted communities decreases. Finally, we introduced here a new kind of data set and make it available to research community.

Future work along these lines might include studies that use these data as an input to propose DTN mobility models, producing interactions between entities similar to the one observed in this paper. Also, these data, in addition to others available from CRAWDAD [3], can be used to studies communities of people. Having the knowledge of such communities or being able to detect them would be of great help to propose efficient communication schemes. Finally, these data can be used as an input to simulators to evaluate protocols designed for DTN scenarios.

Acknowledgements

We would like to thank all the participants of the experiments and we especially acknowledge Pan Hui for his useful support. He helped us to program and package the iMotes, and to extract data from them after the experiment.

7. REFERENCES

- [1] J. Burgess, B. Gallagher, D. Jensen, and B. N. Levine. MaxProp: Routing for vehicle-based disruption tolerant networking. In *Proc. Infocom*, 2006.
- [2] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on the performance of opportunistic forwarding algorithms. In *Proc. INFOCOM*, 2006.
- [3] CRAWDAD a Community Resource for Archiving Wireless Data At Dartmouth. <http://crawdad.cs.dartmouth.edu>.
- [4] Delay Tolerant Network Research Group (DTNRG). <http://www.dtnrg.org>.
- [5] N. Eagle and A. Pentland. Reality mining: Sensing complex social systems. *Personal and Ubiquitous Computing*, 2005.
- [6] T. Henderson, D. Kotz, and I. Abyzov. The changing usage of a mature campus-wide wireless network. In *Proc. Mobicom*, 2004.
- [7] M. Kim and D. Kotz. Classifying the mobility of users and the popularity of access points. In *Proc. LoCA*, 2005.
- [8] J. Lawrence and T. Payne. Exploiting familiar strangers: creating a community content distribution network of co-located individuals. In *Proc. FOAF*, 2004.
- [9] J. Leguay, T. Friedman, and V. Conan. Evaluating mobility pattern space routing for DTNs. In *Proc. INFOCOM*, 2006.
- [10] A. Lindgren, A. Doria, and O. Schelén. Probabilistic routing in intermittently connected networks. In *Proc. (SAPIR 2004)*, 2004.
- [11] E. Paulos and E. Goodman. The familiar stranger: Anxiety, comfort, and play in public places. In *Proc. ACM SIGCHI*, 2004.
- [12] T. Spyropoulos, K. Psounis, and C. Raghavendra. Spray and wait: An efficient routing scheme for intermittently connected mobile networks. In *Proc. WDTN*, 2005.
- [13] C. Tudeuce and T. Gross. A mobility model based on wlan traces and its validation. In *Proc. INFOCOM*, 2005.
- [14] A. Vahdat and D. Becker. Epidemic routing for partially connected ad hoc networks. Technical Report CS-200006, Duke University, April 2000.