Fine-Grained Tracking of Human Mobility in Dense Scenarios

Sabrina Gaito, Elena Pagani, Gian Paolo Rossi
Computer Science Dept., Università degli Studi di Milano
v. Comelico 39, I-20135 Milano, Italy
Email: gaito@dsi.unimi.it, {pagani,rossi}@dico.unimi.it

I. Addressed Problem

This paper envisions an urban scenario where people carry radio devices that can be dynamically networked, by exploiting human contact opportunities, to create unplanned, improvised and localized wireless connectivity, which has been recently called Pocket Switched Networks (PSN). PSNs are a sub-class of Delay Tolerant Networks, but the fascinating idea of considering people as the switching nodes of a mobile network infrastructure is progressively moving the application of DTN from rural, or deep space, to urban scenarios where nodes are more densily distributed, and people contacts are more frequent and dependent on their social relationships. In such a scenario, the design of forwarding protocols is heavily influenced by human mobility, whose understanding has motivated a large amount of researches mainly addressed (i) to the analysis of the distribution of contacts and their duration, (ii) to derive the mobility models, to seamlessly re-create the real patterns of human mobility, and, more recently, (iii) to capture the role of human social attitudes and relationships [5]. This process of understanding and properly reproducing the human mobility is far from fulfilment when densily populated settings (urban locations, workplaces, small campuses) are involved.

The paper focuses on the radio device (*Pocket Mobility Trace Recorder*, or PMTR) we have on purposely designed and developed to improve this understanding by enabling the gathering of rich and detailed mobility data sets from experiments in real mobility settings. PMTRs actually have a couple of novelty points that make them different from devices adopted in similar experiments: firstly, they ensure a very low energy consumption and, secondly, they allow to operate at a high beaconing rate. Low energy consumption allows to perform experiments lasting several consecutive days without the need of battery changes, thus enabling continuous data recording. The high beaconing rate allows to detect contacts lasting a few seconds. The currently available data sets [1] report contacts lasting not less than a few minutes (let us say, 2 or 5 min.), thus missing a huge of shorter contacts whose presence has been proved by our experiments and whose importance can be motivated by observing that they may contribute to heavily reduce the delivery latency and provide a ready-to-use bandwidth to transport short messages.

The main contribution of the paper is twofold: we firstly describe the architecture of the radio devices and, secondly, we provide some evidence of the impact short contacts have on forwarding in dense settings.

It is easy to argue that the ability of detecting short contacts can be constrained by the underlying radio technology, e.g. Bluetooth is not an enabling technology, and can affect the device battery life time. As a consequence, if a forwarding algorithm has been designed to take advantage of short contacts, then the extra power consumption, due to the high beaconing rate, should be compensated by shortening the radio range. We can show how this approach, in dense spaces, has positive effects on the overall system energy consumption without negatively affecting the delivery performances.

II. Pocket Mobility Trace Recorder

The design of a specific device for trace recording is mainly motivated by the need of observing and recording very short contact periods, few seconds, that arise from random mobility in a workplace. As a consequence, PMTRs have been designed to operate with beaconing times ranging from 1 sec. to some configurable value that depends on the mobility environment we wish to observe. Secondly, the devices have to enable unmanned experiments lasting 3-4 weeks without batteries substitution. Whenever a beacon is received from a given encounter, let us say PMTR$_e$, the device creates a new entry in the local *contact_log* if no entry is active for PMTR$_e$. The beacon is discarded otherwise, while the entry is maintained and contains the following items: i) local ID and ID of the encounter node; ii) the time stamp of the first contact; and iii) the time stamp of the contact closing event. An entry in the contact_log is *closed* when the beaconing from the encounter node has been missing for more than t seconds, with $t = 60$ seconds in our experiments. The local memory size should be dimensioned to store the contacts of experiments lasting up to 3-4 weeks. Our test beds have generated on average 2000 contacts per device with beaconing time set to 1 sec. No specific bandwidth and processing requirements have been envisaged for PMTRs.

The PMTR architecture uses the Cypress CY8C29566 micro-controller and the radio module AUREL, model RTX-RTLP. The radio range has been limited to 10 meters in order to reduce the power consumption and to maintain multi-hop paths between end-systems. This combination allows a very low power consumption that lets the experiments last for the

Fig. 1. Circuitry and package of a PMTR. In the packed version, the antenna is fixed all around the box to avoid accidental disconnection.

TABLE I
UNICAST ROUTING MEASURES

	threshold=2'	threshold=5'
# short contacts	46.67%	66.30%
all contacts - capacity	1976 Mb	1976 Mb
long contacts - capacity	3180 Mb	4980 Mb

required time with common batteries NiMh, AA 1.2 V. Each PMTR has a 1 MB flash memory where more than 50K contacts can be stored. The PMTR implements a CSMA non-persistent MAC protocol that regulates access to the 2400 b/s channel. Manchester coding is used at the physical layer. The layer 2 beacons are the only frames that a node has to exchange with its neighbors. The beacon payload is composed of: the PMTR identifier; a set of bits representing the internal state (and which are used for diagnosis purposes); and the current time. The local clock value is set at the configuration time. Each PMTR uses a USB interface to communicate with the Pocket Viewer PC, running the Desktop application software, which has been used to configure the devices, collect the recorded data at the end of the experiment and support data analysis and device monitoring. A PMTR has size $6 \times 10 \times 3$ cm. ($2.36 \times 3.94 \times 1.18$ in., see Fig.1).

This paper reports the data set we obtained from the experiment that was run for 15 working days in November 2008. The PMTRs have been distributed among 16 faculty members, 24 Master and PhD students, 4 persons of the technical staff, and 5 fixed positions (2 coffee shops, a break area, the department entrance and the entrance of the classrooms). Part of the faculty members and of the PhD students (3 persons) have their offices in a different building, which is also used for lessons. At the end of the experiment, 5 devices have been excluded as they showed problems in contact recording, or were not correctly used by their owners. People involved in the experiment have been carefully profiled in terms of mutual friendship relations and movement habits, in order to validate the contacts recorded by their devices.

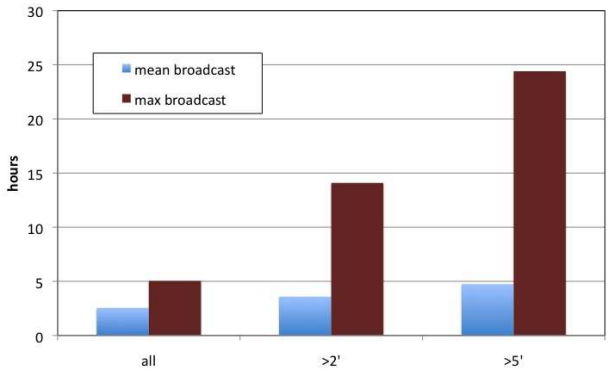

Fig. 2. Latency for broadcast transmissions

III. KEY CONCEPTS AND RESULTS

Our analysis mainly focuses on the influence of short contacts on the opportunistic communications. Contacts lasting less than 5 min. represent the large part of the overall recorded contacts, i.e. 81.55%, due to the combined effects of densily populated spaces and the reduced radio range. To estimate their usefulness for routing, we developed a simple greedy algorithm able to identify the *optimal in latency* path between all source-destination pairs. In particular, we focused on applications where a source needs to send a notification to the devices in the area as soon as the message is generated; both broadcast and unicast schemes are considered. Two performance indexes have been evaluated: the number of short (long) contacts forming the path, and their respective contribution to communication in terms of link capacity. By assuming a 1 Mbps bidirectional bandwidth of wireless links, we obtain capacity estimates by measuring the overall length (in seconds) of the contacts [4]. Two thresholds have been adopted to distinguish short and long contacts, namely, 2 and 5 min.

In Table I, we report the percentage of short and long contacts used by the paths, and the average overall path capacity. In Fig.2, we show the latencies measured for broadcast diffusions. Measures concern three cases: all contacts are used for forwarding, or only long contacts – according to the two thresholds above – are used. The role of short contacts is not negligible, as they represent at least half of the hops along the fastest paths. The capacity of a path is obtained by summing the bandwidth of each contact used, as defined before. By jointly analyzing the latency and the capacity, it can be observed that, in systems where short contacts are not detected, the delivery delay increases and almost doubles if only contacts lasting more than 5 minutes are usable. Unicast communications show a comparable behavior, yielding an average latency that ranges from 40 minutes using all contacts to one hour if only contacts longer than 5 minutes are considered.

As a further measure of the potentiality of this kind of networks, we evaluated the aggregated bandwidth offered by the contacts. Sampling has been performed with 1 sec. gran-

TABLE II
BANDWIDTH AVAILABILITY

	threshold=2'	threshold=5'
$Contrib$	4.56%	9.46%
$Aggbwth_{all}$	7.03 Mbps	7.03 Mbps
$Aggbwth_{long}$	6.70 Mbps	6.36 Mbps

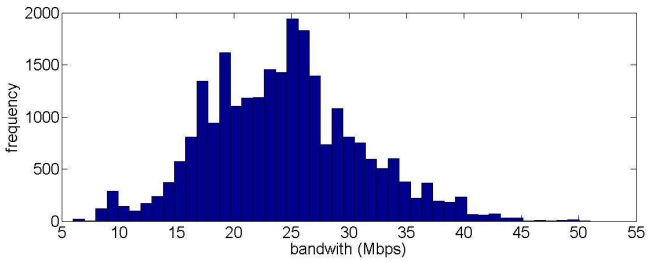

Fig. 3. Histogram of the aggregated bandwidth

ularity, counting the number of links existing in each second, distinguished between short and long links. The maximum number of contacts per sec. is 51, with a mean of 7 and a median of 4; the percentage of empty seconds is 61.79%, accounting for nights and weekends. Two indexes are used to measure the impact of short contacts: let us define $Contrib = \sum lengths\ of\ short\ contacts / \sum lengths\ of\ all\ contacts$; while $Aggbwth_{all}$ and $Aggbwth_{long}$ measure the average aggregated bandwidth on all and on long contacts respectively. In Table II, the two indexes are evaluated over the whole experiment, for the two thresholds above. Short contacts provide up to 10% of the contact time available, and almost 1 Mbps of bandwidth. Hence, they offer a ready-to-use bandwidth with low latency. If users need to exchange large amount of data, in this case they shall use long contacts that are able to offer a higher bandwidth at the expenses of longer delays. In order to know how long a contact could last, a history of measurements of past contacts or the knowledge of user's social relationships could be used [3]. However, from the survey we conducted on almost 300 persons it emerges that users plan to use PSNs mainly to diffuse short notifications, while delaying the sharing of large contents when Internet access is available.

Fig.3 shows the histogram of the aggregated bandwith computed on contacts occurred during the most crowded working day. The bandwith is normally distributed around a mean value of 24.2 Mbps with a standard deviation of 6.8 Mbps. This value changes during the day as shown in Fig.4. It increases in the morning time till an absolute maximum at midday and a secondary maximum in the afternoon. This difference in the number of contacts during the day induces a different use of nodes as relays. In the early morning a crucial node is the doorkeeper since few persons are at office. Thus very short contacts are prevalent in the forwarding path. When people reach their office and the environment becomes denser, messages can be transmitted from office to office and long contacts become important.

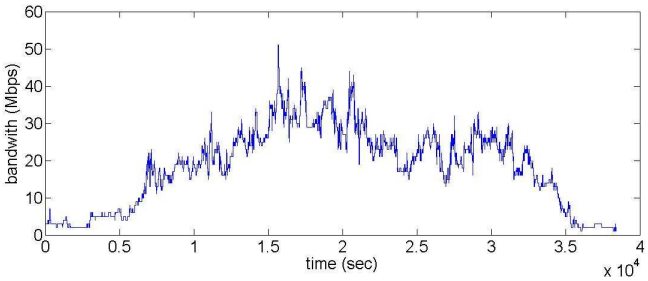

Fig. 4. Behavior of the aggregated bandwidth during a working day

IV. YET ANOTHER NETWORK INFRASTRUCTURE

When considering the application of PSNs in urban dense scenarios, people (and referees, as well) are used to ask why should we introduce a novel, delay-prone network when we have plenty of no-delay infrastructures. We asked the same, without obtaining a definitive answer. However, we surveyed nearly 300 people, among both students and faculty members, and we observed a growing, although still unclearly formulated, interest to address personal communications outside the institutional IT platform because of costs or a feeling of privacy and informality. The growing success of applications such as Twitter (twitter.com), which keeps friends and coworkers frequently connected, and the constant growth of demand for Mobile Chat and Messaging are showing this trend in human interactions and impose to verify if PSNs are suitable to support a part of this type of traffic. While the reader may find in [2] a more detailed analysis of the argument, we think that is not necessary to have the application before the technology. By contrast, we believe that, once deployed, the informal, easy and spontaneous access to this further form of connectivity will incentive the growth of a new family of applications centred on the innovative notion of 'opportunistic social network'.

Acknowledgments: This work has been partially funded by the Italian Ministry of University and Research in the framework of the "Context-Aware RouTing Over Opportunistic Networks (CARTOON)" PRIN Project (http://cartoon.nptlab.dico.unimi.it).

REFERENCES

[1] CRAWDAD Project, http://crawdad.cs.dartmouth.edu/
[2] J. Crowcroft, E. Yoneki, P. Hui, and T. Henderson, *Promoting Tolerance for Delay Tolerant Network Research*. ACM Computer Communication Review, 38(5):63-68, Oct. 2008.
[3] S. Gaito, E. Pagani, and G.P. Rossi, *Opportunistic Forwarding in Workplaces*. To appear in Proc. 2nd ACM SIGCOMM Workshop on Online Social Networks (WOSN), Barcelona (Spain), Aug. 2009.
[4] D. Hay, and P. Giaccone, *Optimal Routing and Scheduling for Deterministic Delay Tolerant Networks*. Proc. IEEE WONS 2009, Snowbird, Utah, USA, Feb. 2009.
[5] Pan Hui, *People are the Network: Experimental Design and Evaluation of Social-based Forwarding Algorithms*. Computer Laboratory Technical Report UCAM-CL-TR-713, Cambridge University, March 2008.