Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms

Augustin Chaintreau‡, Pan Hui*, Jon Crowcroft*, Christophe Diot‡, Richard Gass†, and James Scott†,

‡Thomson Research
46 quai A. Le Gallo
92648 Boulogne FRANCE
augustin.chaintreau@thomson.net
christophe.diot@thomson.net

*University of Cambridge
15 JJ Thomson Avenue
Cambridge, CB3 0FD, UK
pan.hui@cl.cam.ac.uk, jon.crowcroft@cl.cam.ac.uk
richard.gass@intel.com, james.w.scott@intel.com

† Intel Research

Abstract—Studying transfer opportunities between wireless devices carried by humans, we observe that the distribution of the inter-contact time, that is the time gap separating two contacts of the same pair of devices, exhibits a heavy tail such as one of a power law, over a large range of value. This observation is confirmed on six distinct experimental data sets. It is at odds with the exponential decay implied by most mobility models. In this paper, we study how this new characteristic of human mobility impacts a class of previously proposed forwarding algorithms. We use a simplified model based on the renewal theory to study how the parameters of the distribution impact the delay performance of these algorithms. We make recommendation for the design of well founded opportunistic forwarding algorithms, in the context of human carried devices.

I. Introduction

The increasing popularity of devices equipped with wireless network interfaces (such as cell phones or PDAs) offers new communication services opportunities. Such mobile devices can transfer data in two ways - transmitting over a wireless (or wired) network interface, and carrying from location to location by their user (while stored in the device). Communication services that rely on this type of data transfer will strongly depend on human mobility characteristics and on how often such transfer opportunities arise. Therefore, they will require fundamentally different networking protocols than those used in the Internet. Since two (or more) ends of the communication might not be connected simultaneously, it is impossible to maintain routes or to access centralized services such as the DNS.

In order to better understand the constraints of opportunistic data transfer, we take an experimental approach. We analyze six distinct data sets, three of which we have collected ourselves. We define the inter-contact time as the time between two transfer opportunities, for the same devices. We observe in the six traces that the inter-contact time distribution follows a heavy tailed distribution on a large range of values. Inside this range the inter-contact time distribution can be compared to the one of a power-law. We study the impact of those heavy tailed inter-contact times on the actual performance and theoretical limits of a general class of opportunistic forwarding algorithms that we call "naive forwarding algorithms". Algorithms in this class do not use the identity of the devices that are met, nor the recent history of the contacts, or the time of the day, in order to make forwarding decision. Instead forwarding decision are based on forwarding rules statically defined that bound the number of data replicates, or the number of hops.

Based on our experimental observations, we develop a simplified model of opportunistic contact between human-carried wireless devices. It is based on several independence assumptions which are usually met, at least implicitly, in the literature of mobile ad-hoc routing. We do not claim that this model is satisfactory to have accurate performance of different forwarding algorithms. It rather serves our purpose which is to demonstrate how heavy tail inter-contact times influence the performance of naive forwarding algorithms in opportunistic transmission conditions – and how these forwarding algorithms should be configured to offer reasonable performance guarantee.

Our experimental results are presented in Section II. In Section III, we model contact opportunities based on our observations and we analyze the delay that wireless devices would experience using a class of forwarding algorithm previously studied in the literature. Section IV is dedicated to related works. The paper concludes with a brief summary of contributions and presentation of future

work, including a discussion of our assumptions.

II. Experimental analysis

A. Data sets

In order to conduct informed design of forwarding algorithms between devices carried by humans, it is important to study data on the frequency and duration of contacts between them. Ideally, an experiment would cover a large user base over a large time period, as well as include data on connection opportunities encountered twenty-four hours a day, with a granularity measured in seconds.

We examined two types of data sets. First, we used traces made available to us by people who have performed previous measurement exercises. Three data sets emerged, namely from UCSD [1], Dartmouth University [2] and the University of Toronto [3]. We complemented these traces with three of our own experiments. These six experiments use different user populations as well as different wireless technologies. The characteristics of the six data sets, explained below, are shown in Table I.

1) External data sets: UCSD and Dartmouth make use of WiFi networking, with the former including client-based logs of the visibility of access points (APs), while the latter includes SNMP logs from the access points. The durations of the different logs traces are three and four months respectively. Since we required data about device-to-device transmission opportunities, the raw data sets were unsuitable for our experiment and required pre-processing. For both data sets, we made the assumption that mobile devices seeing the same AP would also be able to communicate directly (in ad-hoc mode), and created a list of transmission opportunities by determining, for each pair of devices, the set of time regions for which they shared at least one AP.

Unfortunately, this assumption introduces inaccuracies. On one hand, it is overly optimistic, since two devices attached to the same access point may still be out of range of each other. On the other hand, the data might omit connection opportunities, since two devices may pass each other at a place where there is no instrumented access point, and this contact would not be logged. In addition, the UCSD data set is more exhaustive than the Dartmouth one, since it logs all reachable APs for each client at each time slot, while the Dartmouth data only logs the associated AP. Another issue with these data sets is that the devices are not necessarily co-located with their owner at all times (i.e. they do not always characterize human mobility). Despite these inaccuracies, the WiFi traces are a valuable source of data, since they span many months and include thousands of devices. In addition, considering that two devices, connected to the same AP, are potentially in contact is not altogether unreasonable, as these devices could indeed communicate through the AP, without using end-to-end connectivity.

The University of Toronto have collected traces from 20 Bluetooth-enabled PDAs that were distributed to a group of students. These devices performed a Bluetooth inquiry each 100s and this data was logged. This methodology does not require devices to be in range of any AP in order to collect contacts, but it does requires that the PDAs are carried by subjects and that they have sufficient battery life for them to participate in the data collection. Data may be collected over a long period if devices are recharged. The data set we use comes from an experiment that lasted 16 days.

2) iMote-based experiments: In order to complement the previous traces, we built our own experiment using Intel iMotes, which are embedded devices similar to Crossbow motes[1], but with the key feature (for our experiments) that they communicate via Bluetooth. We programmed the iMotes to log contact data for all visible Bluetooth devices (including iMotes as well as other Bluetooth devices such as cell phones). Each contact is represented by a tuple (MAC address, start time, end time). The experimental settings are described in detail in [4]; an anonymized version of our data will be made available to other research groups on demand.

Three iMote-based experiments were conducted. The first included eight researchers and interns working at Intel Research in Cambridge. The second obtained data from twelve doctoral students and faculty comprising a research group at the University of Cambridge Computer Lab. The third experiment was conducted during the IEEE INFOCOM 2005 conference in Miami where 41 iMotes where carried by attendees for 3 to 4 days. iMotes contacts were classified into two groups: iMotes recording the sightings of another iMotes are classified as "internal" contacts, while sightings of other types of Bluetooth devices are called "external" contacts. The external contacts are numerous and include anyone who has an active Bluetooth device in the vicinity of the iMote carriers, thereby providing a measure of actual wireless networking opportunities present at that time. The internal contacts, on the other hand, represent the data transfer opportunities among participants, if they were all equipped with devices which are always-on and always-carried.

---

[1]www.xbow.com

User Population	Intel	Cambridge	Infocom	Toronto	UCSD	Dartmouth
Device	iMote	iMote	iMote	PDA	PDA	Laptop/PDA
Network type	Bluetooth	Bluetooth	Bluetooth	Bluetooth	WiFi	WiFi
Duration (days)	3	5	3	16	77	114
Granularity (seconds)	120	120	120	120	120	300
Devices participating	8	12	41	23	273	6648
Number of internal contacts	1,091	4,229	22,459	2,802	195,364	4,058,284
Average # Contacts/pair/day	6.5	6.4	4.6	0.35	0.034	0.00080
Recorded external devices	92	159	197	N/A	N/A	N/A
Number of external contacts	1,173	2,507	5,791	N/A	N/A	N/A

TABLE I

COMPARISON OF DATA COLLECTED IN THE SIX EXPERIMENTS.

B. Definitions

We are interested in how the characteristics of transfer opportunities impact data forwarding decisions. In this paper, we focus on how often such opportunities occur. We decided not to attempt to analyze how much data can be transported for each of them, as this strongly depends on factors such as the transmission protocol, the antennas used, and other factors that could be modified to provide improved transmission performance. In our analysis in Section III, we address two extreme cases corresponding to a lower and upper bounds of the amount of data that may be transferred in each connection opportunity.

We define the *inter-contact time* as the time elapsed between two successive contacts of the same devices. Inter-contact time characterizes the frequency with which packets can be transferred between networked devices; it has rarely been studied in the literature. Two remarks must be made at this point:

First, the inter-contact time is computed once at the end of each contact period, as the time interval between the end of this contact and the next contact with the same devices[2]. Another option would be to compute the remaining inter-contact time seen at any time, i.e at time t, for each pair of devices: the remaining inter-contact time is the time it takes after t, before a given pair of devices met again (a formal definition is given in Section III). Inter-contact time and remaining inter-contact time have different distributions, which are related, for a renewal process, via a classical result known as the waiting time paradox, or inspection paradox (see p.147 in [5]). A similar relation holds for stationary process, in the theory of Palm Calculus (see p.15 in [6]). We choose to study the first definition of "inter-contact time seen at the end of a contact period", as the second gives too much weight to large values of inter-contact times. In other words the definition that was chosen is the most

[2]Inter-contact starting after the last contacts recorded for this pair of devices were not included.

conservative one in the presence of large values.

Second, the inter-contact time distribution is influenced by the duration and the granularity of the experiment. Inter-contact times that last more than the duration of the experiment cannot be observed, and inter-contact times close to the duration are less likely to be observed. In a similar way, inter-contact times that last less than the granularity of the measurement (which ranges from two to five minutes among different experiments) cannot be observed.

Another measure of the frequency of transfer opportunities, that could be considered, is the inter-any-contact, i.e. for a given device, the time elapsed between two successive contacts with any other device. This measure is very much dependent on the deployment of wireless devices and their density during the experiment, as it characterizes time that devices spend without meeting any other device. This measure was studied for most of these datasets in [4], we do not present further results here, due to a lack of space.

C. Inter-contact time characterization

We plot the inter-contact time distribution for all six experiments in Figure 1-2. For the two first iMote experiments (labeled Intel and Cambridge) the distribution of inter-contact were computed using all pairs of two iMotes. Therefore it contains only values associated with internal contacts (However, we observe the exact same properties for external contacts). Data from Toronto experiment were also collected between pairs of experimental Bluetooth devices. Distributions for these three datasets are plotted in Figure 1 (left). Distributions belonging to the iMote based experiment at Infocom is shown in Figure 1 (right), where inter-contacts belonging to both internal and external contacts have been plotted separately for comparison. Figure 2 presents the distribution of inter-contact computed using traces from WiFi experiments. All plots describe the tail distribution function, in log-log scale.
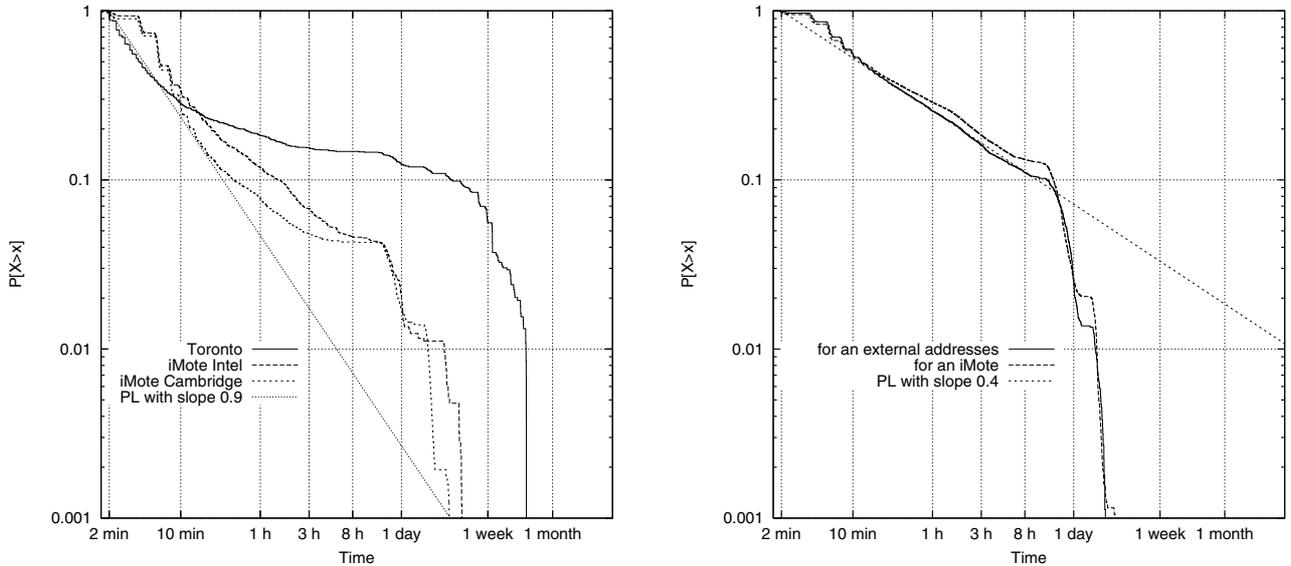
Fig. 1. Tail distribution function of the inter-contact time in six experiments: iMote-based experiment at Intel and Cambridge, and Toronto experiment (left), iMote-based experiment at INFOCOM (right).

The most interesting region is the middle of the graphs, as the leftmost and rightmost parts show artifacts due to the granularity and duration of the experiments as explained above. In this region, all six distributions show the same characteristics: they exhibit an *heavy tail*, that can be approximated or lower bounded by the tail of a power law, over a large range of value. This common property is rather surprising given the diversity of the six data sets. The most notable difference is that the match with a power law, as evidenced by the straightness of the curve, is better for the data sets that are shown in Figures 1 (right) and 2, which contain the largest number of contacts. Figure 1 (right) proves that the distribution is almost unchanged if one consider internal or external contacts. The same results was shown for the two other iMote experiments and are presented in [7].

A power law is characterized by its coefficient reflecting the slope of the line on log-log graphs; we show later that this coefficient is critical for the performance of the forwarding algorithms presented in Section III. For the iMote-based experiments at Intel and Cambridge, and the data collected in Toronto, the tail is lower bounded by a power law with coefficient 0.9 for the range [2 min; 1 day]. The distribution for the iMote-based experiment at Infocom is remarkably close to a power law with coefficient 0.4 on the range [2 min, 16h]. The tail from Dartmouth data can be approximated by a power law with a coefficient of 0.3 on the range [10 min; 1 week]. The tail from UCSD data can be also compared with a power law with coefficient 0.3, but over a more limited range [10 min; 1 day].
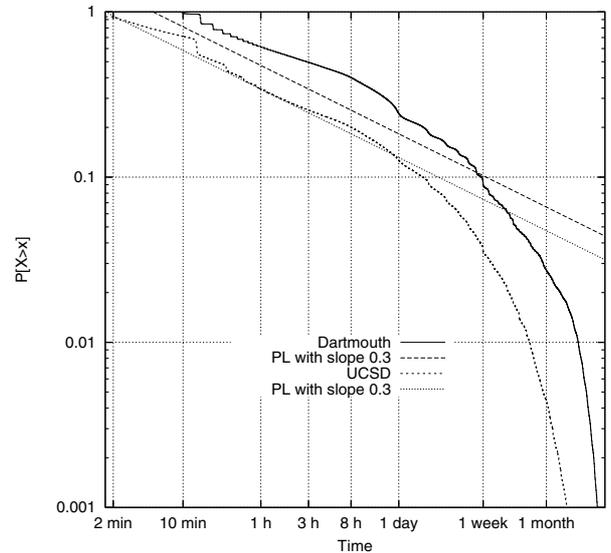

Fig. 2. Tail distribution function of the inter-contact time: data collected in Dartmouth and UCSD.

A tail that can be compared or lower bounded by a power law means the tail distribution function decreases slowly over this range. This contradicts the exponential decay that is implied by many mobility models in the literature (see Section IV). As a result, opportunistic networking algorithms which have been designed around exponential models must be re-evaluated in the light of our observations (see next section). In the iMote-based traces, 8 to 25% of inter-contact times are greater than one hour, and 2 to 3% are greater than one day. In

the Toronto trace, 13% last more than a day, and 7% last more than one week. Similarly in the Dartmouth trace, we find that large inter-contact times are far from negligible: 20% last more than a day, 10% last more than a week. In the UCSD trace, 15% last more than a day, and 4% last more than one week.

While the WiFi experiments have longer durations, longer inter-contact times may be affected by the more limited mobility of laptops or PDAs as their users may not carry them all the time. However, this is a characteristic of how users of wireless devices behave and this should be taken into consideration in the design of forwarding mechanisms. What remains is that the same pattern – a heavy tail that can be compared to the one of a power law – seems to apply to all experiments despite the fundamental differences in methodology and in experimental environments.

III. FORWARDING WITH POWER LAW–BASED OPPORTUNITIES

We now study analytically the impact of our findings on the performance of a class of forwarding algorithms. We define first our model of the opportunistic behavior of mobile users that is based on our experimental observations.

A. Assumptions and Forwarding Algorithms

1) Contact process model: We consider a slotted time $t = 0, 1, \ldots$. For a given pair of devices (d, d'), let us introduce its contact process $(U_t^{(d,d')})_{t \geq 0}$ defined by:

$$U_t^{(d,d')} = \begin{cases} 1 & \text{if } d \text{ and } d' \text{ are in contact during slot } t, \\ 0 & \text{otherwise.} \end{cases}$$

For the pair (d, d') we consider the sequence of the time slots $T_0^{(d,d')} < T_1^{(d,d')} < \ldots < T_k^{(d,d')} < \ldots$ that describes all the values of $t \in \mathbb{N}$ such that $U_t^{(d,d')} = 1$.

We do not include in this model the contact time representing the duration of each contact, assuming that each contact starts and ends during the same time slot. This is justified here by the fact that we are interested in a model accounting for consequences of large values of the inter-contact time. It was observed (see [7]) that the contact time distribution is also heavy tailed, but it takes smaller values, by several orders of magnitude, than the ones of the inter-contact time.

Under this condition, the time $\tau_k^{(d,d')} = T_{k+1}^{(d,d')} - T_k^{(d,d')}$ for any d, d' and $k \geq 0$ is the inter-contact time after k_th contact of this pair. We suppose in our model that it is distributed according to the law X, that is a power law distribution with coefficient α:

$$P[X \geq t] = t^{-\alpha} \quad \text{for all } t = 1, 2, \ldots \quad . \tag{1}$$

In particular, variable X is not bounded but it is almost surely finite.

In addition we assume that the contact process $(U_t^{(d,d')})_{t \geq 0}$ of each pair is a renewal process, and that contact processes associated with different pairs are independent. In other words, the inter-contact times in the sequence $(\tau_k^{(d,d')})_{k \geq 0}$ are i.i.d., for all (d, d'), and sequences belonging to different pairs are independent.

We come back to these assumptions later in Section V. Note that these assumptions are verified, or implicitly assumed, in most of the analysis of currently proposed mobility models. This is because it is typically very difficult to analyze models where dependence may arise between different devices or between successive events occurring with one or more devices.

Even if we do not explicitly model the contact time (each contact lasts one time slot), we need to take into consideration the fact that a contact may last long enough to transmit a significant amount of data. We then introduce two situations:

- the *short contact case* : where only a single data unit of a given size can be sent between devices in a single time slot where they are in contact.
- the *long contact case* : where two devices in contact can exchange an arbitrary amount of data during a single time slot.

These two cases represent a lower and an upper bounds for the evaluation of bandwidth. The number of data unit transmitted in a contact (whether short or long) is defined as a data bundle. the long and the short case differ from a queuing standpoint. In the long contact case, the queue is emptied any time a destination is met. In the short contact case, only one data unit is sent and therefore, data can accumulate in the memory of the relay device.

Note that our model is not taking into account explicit geographical locations and movement of devices, as it assumes directly that the processes of contacts between them are given. The results of this section extend to any mobility model which creates independent contact processes for all pairs of devices, that follow this same law.

Before describing forwarding algorithms that we consider, let us introduce for any pair of device (d, d'), the **remaining inter-contact time**, observed at time slot t: It is denoted by $R_t^{(d,d')}$ and defined as

$$R_t^{(d,d')} = \min \left\{ t' - t \mid t' \geq t \text{ and } U_{t'}^{(d,d')} = 1 \right\} \ .$$

2) Forwarding algorithms: We are interested in a general class of forwarding algorithms, which all rely on other devices to act as relays, carrying data between

a source device and a destination device that might not be contemporaneously connected. These relay devices are chosen purely based on contact opportunism and not using any stored information that describes the current state of the network. The only information used in forwarding is the identity of the destination so that a device knows when it meets the destination of a bundle. We call such algorithms "naive", although they could be in reality quite complex and, as we will see, very efficient in some cases.

The following two algorithms provide bounds for the class of algorithm described above:

- wait-and-forward: The source waits until its next direct contact with the destination to communicate.
- flooding: a device forwards all its received data to any device which it encounters, keeping copies for itself.

The first algorithm uses minimal resources but can incur very long delays and does not take full advantage of the ad-hoc network capacity. The second algorithm, that was initially proposed in [8], delivers data with the minimum possible latency, but does not scale well in terms of bandwidth, storage, and battery usage. In between these two extreme algorithms, there is a whole class of algorithms that play on the number of relay devices to maximize the chance of reaching the destination in a bounded delay while avoiding flooding. The most important reason not to flood is to minimize memory requirements and related power consumption in relay devices, and to delete the backlog of previously sent message that are still waiting to be delivered, and could be outdated. Some strategies, based on time-outs, buffer management, limit on the number of hops and/or duplicate copies have been proposed (see [8], [9], [10]) to minimize replication and backlog.

B. Analysis of the two-hop relaying algorithm

Having described the class of "naive" algorithms we are considering in this work, we now introduce the two-hop relaying algorithm [11], and evaluates its performance for the model of power law inter-contacts that we have described. Results are generalized to the class of naive algorithms in the following section.

1) Description: The two-hop relaying algorithm was introduced by Grossglauser and Tse in [11]. This forwarding algorithm operates as follow: when a source has a bundle to send to a destination, it forwards it once to the first devices that it meets. This first device is either a relay device or the destination itself. If it is the destination, the bundle is delivered in one hop; otherwise the device acts as a relay and stores the bundle

in a queue corresponding to this destination. Bundles from this queue will be delivered when the relay device meets the destination. Bundles for the same destination are delivered by a relay device in a first-come-first-served order. As queuing may occur in the devices that act as relays, in the short contact case, the forwarding process of bundles sent by the source to a relay needs to be of lower intensity than the bundles sent by this relay to the destination. This is the case in the implementation proposed in [11] and we make the same assumption below.

We choose this algorithm to start our study of the impact of power law inter-contact times on opportunistic forwarding for the following three reasons:

- In the short contact case, this algorithm was shown to maximize the capacity of dense mobile ad-hoc networks, under the condition that devices locations are i.i.d., distributed uniformly in a bounded region.
- This result depends strongly on the mobility process of devices. Authors of [11] assumed an exponential decay of the inter-contact time. The same result has been proven for devices following for instance the random way-point mobility model [12].
- [11] and [12] have shown that that data experiences a finite expected delay under these conditions.

2) Analysis: We consider N mobile devices which transmit data according to the two-hop relaying algorithm described above. Instead of the mobility model used in [11] we assume that contacts between devices follow the model that we have introduced in the beginning of this section.

To ensure stability in the relay's queuing mechanism, we assume that the source s is not saturated: bundles are created at s during a sequence of time slots. The same assumption is made for the long contact case although stability of the queue occupancy is not an issue in this context as the queue is emptied after each contact with the destination.

We have, as a consequence from the regenerative theorem (or Smith's formula).

Theorem 1 *For a pair of source-destination devices* (s, d)*, let* $t_k^{(s)}$ *be the time when the k_th bundle is created at s to be sent to d, and let* $t_k^{(d)}$ *be the time when it is delivered to d. We have for* $D_k = t_k^{(d)} - t_k^{(s)}$*:*

(i) If $\alpha < 2$*,* $\lim_{k\to\infty} \mathbb{E}[D_k] = +\infty$*.*

(ii) If $\alpha > 2$*, and we assume that all contacts are long,* $\lim_{k\to\infty} \mathbb{E}[D_k] = \bar{D} < +\infty$ *and we have*
$$\bar{R} \leq \bar{D} \leq 2\bar{R} \quad where \quad \bar{R} = \frac{1}{2} + \frac{\mathbb{E}[X^2]}{2.\mathbb{E}[X]} .$$

(iii) If $\alpha > 3$*, and we assume that all contacts are short, for a stationary point process* $(t_k^{(s)})_{k\geq 0}$ *with*

intensity $\lambda < (N-1)/\bar{R}$, there exists a stationary regime where the delay of a bundle has a finite expected value $\bar{D}$ verifying for a constant σ_R

$$\frac{1}{2}\left(1 + \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]}\right) \leq \bar{D} \leq \frac{\sigma_R}{\frac{N-1}{\lambda} - \bar{R}} \ .$$

Proof: We study first the case of long contacts, where any amount of information may be exchanged when a contact occurs between two devices.

We analyzed here a single source-destination pair. The two-hop relaying strategy uses multiple routes to transport bundles belonging to this pair; that is because any other contacted device may act as a relay. Let us denote by $t_k^{(s)}$ the time when the k_th bundle is created in the source for this destination. This bundle is transmitted to the first relay that is met by s after time $t_k^{(s)}$. The relay chosen is $r_k = \text{argmin}_{r' \neq s} R_{t_k^{(s)}}^{(s,r')}$; and this transmission occurs at time $t_k^{(r)} = t_k^{(s)} + \min_{r' \neq s} R_{t_k^{(s)}}^{(s,r')}$. The bundle is then delivered to destination d at time $t_k^{(d)} = t_k^{(r)} + R_{t_k^{(r)}}^{(r_k,d)}$. Note that in this case:

$$D_k = \min_{r \neq s} R_{t_k^{(s)}}^{(s,r)} + R_{t_k^{(r)}}^{(r_k,d)} \ . \tag{2}$$

Let us first establish the positive result (ii) that the two-hop relaying strategy achieves a delay with finite mean if $\alpha > 2$.

Proving (ii) **:** In this case, $\mathbb{E}[X^2]$ is finite, and $\mathbb{E}\left[\sum_{t=T_0^{(d,d')}}^{T_1^{(d,d')}-1} R_t^{(d,d')}\right] = \mathbb{E}[X(X+1)/2] < \infty$, for any pair (d,d') of devices. By Smith's formula (see (5) in the appendix), we have $\lim_{t\to\infty} \mathbb{E}\left[R_t^{(d,d')}\right] = \frac{\mathbb{E}[X^2]+\mathbb{E}[X]}{2\mathbb{E}[X]}$.

The process $(\min_{r\neq s} R_t^{(s,r)})_{t\geq 0}$ is taken as a minimum of a finite number of independent processes, corresponding to pairs $\{(s,r) \,|\, r \neq s\}$, which all have the same law.

Hence, $\quad \lim_{t\to\infty} \mathbb{E}\left[\min_{r\neq s} R_t^{(s,r)}\right] \leq \frac{\mathbb{E}[X^2] + \mathbb{E}[X]}{2\mathbb{E}[X]} \ .$

Lemma 2 can then be applied to this process, with $(t_k^{(s)})_{k\geq 0}$ which is independent from it; this proves $\lim_{k\to\infty} \mathbb{E}\left[\min_{r\neq s} R_{t_k^{(s)}}^{(s,r)}\right] \leq \frac{\mathbb{E}[X^2]+\mathbb{E}[X]}{2\mathbb{E}[X]} \ .$

If we consider the collection of random variables $((R_t^{(r,d)})_{t\geq 0})_{r\neq s}$, the condition (i) of Lemma 2 is met. As $(t_k^{(r)})_{k\geq 0}$ and $(r_k)_{k\geq 0}$ depend only on $(t_k^{(s)})_{k\geq 0}$ and contacts processes belonging to other pairs than $\{(r,d) \,|\, r \neq s\}$, they are independent from the collection above, and we have $\lim_{k\to\infty} \mathbb{E}\left[R_{t_k^{(r)}}^{(r_k,d)}\right] = \left(\frac{1+\mathbb{E}[X^2]}{2\mathbb{E}[X]}\right) \ .$ Using (2), we have

$$R_{t_k^{(r)}}^{(r_k,d)} \leq \ D_k = \min_{r\neq s} R_{t_k^{(s)}}^{(s,r)} + R_{t_k^{(r)}}^{(r_k,d)}, \ \text{ and}$$

$$\frac{1}{2}(1 + \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]}) \leq \ \lim_{k\to\infty} \mathbb{E}[D_k] \leq (1 + \frac{\mathbb{E}[X^2]}{\mathbb{E}[X]}) \ .$$

Note that this result holds if the law of X is replaced by any law that admits a finite second moment.

Proving (i)**, for** $1 < \alpha < 2$ **:** As $\alpha > 1$, Smith's Formula (5) holds in this case for any function f verifying the integrability condition.

Let r denote any device different from s. For convenience, let us denote $X_1 = T_1^{(r,d)} - T_0^{(r,d)}$, we have for any A, that may be chosen arbitrary large:

$$\frac{A(A+1)}{2}\mathbb{I}_{\{X_1 \geq A\}} \leq \sum_{t=T_0^{(r,d)}}^{T_1^{(r,d)}-1} \min(R_t^{(r,d)}, A) \leq A.X_1 \ .$$

These variables are positive; they all have a finite expectation by comparison with the right term. This proves the integrability condition required in (5) for the function $f(x) = \min(x, A)$, hence we obtain

$$\lim_{t\to\infty} \mathbb{E}\left[\min(R_t^{(r,d)}, A)\right] \geq \ \frac{\frac{A(A+1)}{2}\mathbb{P}[X_1 \geq A]}{\mathbb{E}[X_1]} \geq \frac{A^2.A^{-\alpha}}{2.\mathbb{E}[X_1]} \ .$$

As this inequality holds for A arbitrary large, and $\alpha < 2$, we have: $\lim_{t\to\infty} \mathbb{E}\left[R_t^{(r,d)}\right] = +\infty$. The collection of processes $((R^{(r,d)})_{t\geq 0})_{r\neq s}$ verifies condition (b) of Lemma 2. As $(t_k^{(r)})_{k\geq 0}$ and $(r_k)_{k\geq 0}$ are independent of this collection, we can therefore deduce that

$$\lim_{k\to\infty} \mathbb{E}\left[R_{t_k^{(r)}}^{(r_k,d)}\right] = +\infty \ \text{ hence } \ \lim_{k\to\infty} \mathbb{E}[D_k] = +\infty \ .$$

Proving (i)**, for** $\alpha \leq 1$ **:** In this case, for any device r, the Markov chain defining $(R_t^{(r,d)})_{t\geq 1}$ is recurrent null, so that Orey's theorem (see [5] p.131) tells us :

$$\lim_{t\to\infty} \mathbb{P}\left[R_t^{(r,d)} = i\right] = 0 \ \text{ for all } i$$

In particular, for any A arbitrary large,

$$\lim_{t\to\infty} \mathbb{P}\left[R_t^{(r,d)} < A\right] = 0 \ \text{ and } \ \lim_{t\to\infty} \mathbb{P}\left[R_t^{(r,d)} \geq A\right] = 1 \ .$$

We have, $\mathbb{E}\left[R_t^{(r,d)}\right] \geq A.\mathbb{P}\left[R_t^{(r,d)} \geq A\right]$. As a consequence, and because the result holds for any arbitrary A, we have $\lim_{t\to\infty} \mathbb{E}\left[R_t^{(r,d)}\right] = +\infty$. This holds for any device r. Another application of Lemma 2 with condition (b) allows us to prove $\lim_{k\to\infty} \mathbb{E}\left[R_{t_k^{(r)}}^{(r_k,d)}\right] = +\infty$.

The result (i) for short contacts can then be deduced as the delay in this case, that may include some queuing, is always greater.

Proving (iii) **:** This result is in fact an extension of a method proved in [13], where the two-hop relaying strategy was analyzed for a mobility model assuming brownian motion. The argument is the following:

Let us focus on the queue belonging to a source-destination pair, in a given relay. We denote arrival of data bundles in this queue by t_k. We introduce W_k that is the "remaining load" in the queue when the bundle

k arrives. W_k is equal to the time needed to transmit to d all the data present in the queue when k arrives. The sequence (W_k) follows the following recurrence equation: $W_{k+1} = (W_k + s_k - (t_{k+1} - t_k))^+$, where s_k denotes the additional time, that is added to W_k, to deduce the time at which packet k leaves this queue. This equation is exactly one of a single server queue of customer arriving at time $(t_k)_{k\geq 0}$, requesting service $(s_k)_{k\geq 0}$. The difficulty comes from the fact that $(s_k)_{k\geq 0}$ is not an i.i.d. sequence as it depends on the value of $(W_k)_{k\geq 0}$: In fact if $W_k = 0$, s_k is equal to $R_{t_k}^{r,d}$. Otherwise, the time when all bundles until $k-1$ are delivered is a time of contact $T_j^{r,d} \geq t_k$ for a certain j, and s_k corresponds to an additionnal inter-contact time $T_{j+1}^{r,d} - T_j^{r,d}$, that is independent from the rest, and follows law X. The key argument that we use here was first proposed in [12]: the law of X is stochastically smaller than the one of $R_t^{r,d}$, hence we can show that the sequence $(s_k)_{k\geq 1}$ is stochastically smaller than the sequence $\tilde{R}_k$, where $\tilde{R}_{k\geq 0}$ is an i.i.d. sequence with the distribution R, that is the one of $R_t^{r,d}$ in steady state.

W is a monotone function of the values of the sequence s_k, hence we can show that the following sequence $(\tilde{W}_k)_{k\geq 0}$, defined by recurrence as $\tilde{W}_k = \left(W_{k-1} + \tilde{R}_{k-1} - (t_k - t_{k-1}) \right)^+$, is stochastically greater than $(W_k)_{k\geq 0}$.

This allows us to prove the stability of the queue if the arrivals in the queue follow a process with intensity $\lambda' \leq \frac{1}{\mathbb{E}[R]}$. As we have assumed $\alpha > 3$, the second moment of the law of R is finite (as can be seen from the expression of π). Once can then use Kingman's bound (see (9) in [14]) to show that under this stability condition, the expected value of W_k verifies: $\mathbb{E}[W_k] \leq \frac{2\sigma_R}{2(1/\lambda' - \mathbb{E}[R])}$. We can apply the same proof to all the $N-1$ nodes in the network, so that the intensity of the creation of data bundle in the source can be $\lambda(N-1)$. ∎

To summarize, we have identified two regions where the behavior of the two-hop relaying algorithm would differ, under the power law inter-contact time assumption: For a value of α that is greater than 2 in the long contact case and 3 in the short contact case, the algorithm converges to a finite expected delay, as in the case of an exponential decay. By opposition For α smaller than 2, the two-hop forwarding algorithm will not converge to a finite expected delay, as the delay that can be expected grows without bound with time. This remains true even for long contact case, where data exchange are unlimited during contacts, and queuing in relay devices have therefore no impact on the delay experienced. In other words, the region $\alpha > 2$ ($\alpha > 3$ in the short contact case) may be thought as the stability region of the two-hop relaying algorithm.

C. Generalization

In this section we characterize the region of stability (defined as the value of α for which an algorithm achieves a bounded delay) for the general class of naive algorithms. We conduct the following proof in the long contact case only.

To do so, we generalize the two-hop relaying algorithm as follows. Instead of sending a single copy of a given data unit to a unique relay, the source will send m copies of each data unit: one to each of the first m relays that it meets. As we have assumed that the contacts processes belonging to these relays are independent, the source may reduce the total transmission delay by increasing its probability to pick a relay with a small delay to the destination among the m relays to which it has forwarded the message. This observation is made rigorous in following lemma:

Lemma 1 Let $(R_t^{(d_1,d_1')})_{t\geq 0}, \ldots, (R_t^{(d_m,d_m')})_{t\geq 0}$ be remaining inter-contact times for m different pairs of devices $(d_i, d_i')_{1\leq i\leq m}$.

We suppose $m > 1$ and that $1 + \frac{1}{m} < \alpha < 2$,

then $\mathbb{E}\left[R_t^{(d_1,d_1')} \right] = \ldots = \mathbb{E}\left[R_t^{(d_m,d_m')} \right] = +\infty$

and $\mathbb{E}\left[\min(R_t^{(d_1,d_1')}, \ldots, R_t^{(d_m,d_m')}) \right] < \infty$.

Proof: To illustrate the proof with simple argument, we treat the case where $m = 2$. We suppose $\frac{3}{2} < \alpha < 2$ and we prove for any two pairs of devices (d, d') and (e, e') that $\mathbb{E}\left[\min(R_t^{(d,d')}, R_t^{(e,e')}) \right] < \infty$. The general case is treated in Appendix B. We decided not to include it directly in the text, as it involves many additionnal notation, with almost the same technique.

As $\alpha > 1$, Lemma 3 (ii) can be applied. The product chain $(R_t^{(d,d')}, R_t^{(e,e')})_{t\geq 0}$ then admits the following stationary distribution: $\pi(i,j) = \frac{(i+1)^{-\alpha}(j+1)^{-\alpha}}{(c_1)^2}$ where $c_1 = \sum_{i'\geq 0}(i'+1)^{-\alpha}$.

$$\mathbb{E}\left[\min(R_t^{(d,d')}, R_t^{(e,e')}) \right] = \frac{1}{(c_1)^2} \sum_{i,j} \frac{\min(i,j)}{(i+1)^{\alpha}(j+1)^{\alpha}}$$

We have, by symmetry

$$\sum_{i,j\geq 1} \frac{\min(i,j)}{i^{\alpha}j^{\alpha}} = 2.\sum_{i,j\geq 1} \frac{\min(i,j)}{i^{\alpha}j^{\alpha}} \mathbb{I}_{\{i\leq j\}}$$
$$= 2.\sum_{j\geq 1} \frac{1}{j^{\alpha}} \left(\sum_{i=1}^{j} i^{1-\alpha} \right)$$

The function $x \mapsto x^{1-\alpha}$ is non-increasing on $]0; +\infty[$, as $\alpha > 1$, hence we have for any $i \geq 1$:

$$i^{1-\alpha} \leq \int_{i-1}^{i} x^{1-\alpha}dx < +\infty \text{ as } \alpha < 2, \text{ hence}$$

$$\sum_{i=1}^{j} i^{1-\alpha} \le \int_0^j x^{1-\alpha} dx = \frac{j^{2-\alpha}}{2-\alpha} \; .$$

$$\sum_{i,j\ge 1} \frac{\min(i,j)}{i^\alpha j^\alpha} \le 2.\sum_{j\ge 1} \frac{1}{j^\alpha}\left(\frac{1}{2-\alpha} j^{2-\alpha}\right)$$
$$\le \frac{2}{2-\alpha}\left(\sum_{j\ge 1} j^{2-2.\alpha}\right) \; .$$

This proves that the expectation is finite if $\alpha > \frac{3}{2}$. ■

This result shows that for α smaller than 2, the expected time to meet the destination is infinite. However, the expected time for the destination to meet a group of m devices may have a finite expected value, provided that $\alpha > 1$ and that m is large enough. This observation is the key component in the next result, which proves that using a two-hop relaying strategy with m relays is sufficient to extend the stability region to any value of $\alpha > 1$. This theorem also proves that the case $\alpha < 1$, which is observed in most data sets, is of a quite different nature, as even unlimited flooding does not achieve a bounded delay. We comment on this difference further in Section V.

Theorem 2 *Let us consider a source destination pair (s,d) and $t_k^{(s)}, t_k^{(d)}, D_k$ defined as in Theorem 1. We assume that all contacts are long.*

 (i) *if $\alpha > 2$, there exists a forwarding algorithm using only one copy of the data, with a finite expected delay. $\lim_{k\to\infty} \mathbb{E}[D_k] = \bar{D} < +\infty$.*

 (ii) *if $1 < \alpha < 2$, $m \in \mathbb{N}$ is chosen such that $\alpha > 1 + \frac{1}{m}$, and the network contains at least $N \ge 2m$ devices, there exists an algorithm using m relay devices such that: $\lim_{k\to\infty} \mathbb{E}[D_k] = \bar{D} < +\infty$.*

 (iii) *if $\alpha \le 1$, for a network containing a finite number of devices, and any forwarding algorithm, including flooding, we have $\lim_{k\to\infty} \mathbb{E}[D_k] = +\infty$.*

Proof: **Proving** (i) is just a reminder of the result of Theorem 1. The two hop relaying algorithm may be chosen and it achieves a finite expected delay.

Proving (ii)**:** Let us assume that $\alpha > 1 + 1/m$ and $N \ge 2m$, where $m \in \mathbb{N}$. The forwarding algorithm that we consider in this case is a two-hop relaying algorithm using m different relays.

STEP 1 : A bundle is created at time t in the source (denoted as device s). It is first transmitted to the m first devices that are met. We estimate first the time when each of these m relays are all contacted and have received the bundle. Let us consider the collection of remaining inter-contact time with all the other devices $(R_t^{(s,r)})_{r\ne s}$. This collection contains $N-1$ variables. If we consider a version of this collection, sorted for each time t, in the increasing order, the time to contact m different devices at time t is the $m\_$th value of this

sorted sequence. Corollary 1, which is a simple variation of Lemma 1 shown in Appendix B, tells that this variable is of finite expected value if $\alpha > 1 + 1/(N-1-m+1)$. This last assumption is automatically verified as $N-1-m+1 = N-m \ge m$ by assumption.

STEP 2 : At time t', a copy of the bundle is present in each of the m relays, that we denote $r_1, \ldots, r_m$. We now consider the vector $(R_{t'}^{(r_1,d)}, \ldots, R_{t'}^{(r_m,d)})$ which describes the times needed for each of this relay to get in contact with the destination. The time length elapsed until the packet is delivered to the destination is taken as the minimum of this values. An application of Lemma 1 tells us that this time is finite expected value.

As a consequence the overall delay, from the time of creation of the bundle in the source, to the delivery at the destination, is the sum of two variables with finite expectations. It is hence of finite expected value.

Proving (iii)**:** Let us consider in this case, for a source s and any other device r in the network, the remaining time $R_t^{(s,r)}$ a time t until the next contact. As $\alpha < 1$, all of this sequences of random variables are irreducible null recurrent Markov chains. By Orey's theorem ([5] p.131), we then have that $\lim \mathbb{P}\left[R_t^{(s,r)} = i\right] = 0$ for all i when t tends to infinity. In particular for any A arbitrary large, we have $\lim \mathbb{P}\left[R_t^{(d,d')} \le A\right] = 0$, so that

$$\mathbb{P}\left[\min_{r\ne s} R_t^{(s,r)} \ge A\right] = \mathbb{P}\left[\bigcap_{r\ne s}\{R_t^{(s,r)} \ge A\}\right] \to_{t\to\infty} 1 \; .$$

Consequently, $\mathbb{E}\left[\min_{r\ne s} R_t^{(s,r)}\right]$ diverges for large t.

As a consequence, starting from any initial condition, the time for a source to reach any other device is of infinite expectation as times increases. No forwarding algorithm, no matter how redundant, can then transport a packet within a finite expected delay, using only opportunistic contact between devices. ■

Note: By comparison, the result (iii) applies to any case that includes short contacts as well as long contacts. A network containing N devices admits forwarding algorithms that achieve a bounded expected delay for any $\alpha > 1 + \frac{1}{\lfloor N/2 \rfloor}$: flooding (that may use up to $N-2$ relays) is one of these, but it not the only one, as a forwarding algorithm using only $\lfloor N/2 \rfloor$ relays is sufficient.

D. Summary, Discussion

At this stage, we have established the following results for the class of so-called naive forwarding algorithm defined in III-A, in the long contact case :

 • For $\alpha > 2$ any algorithm from the class we considered achieve a delay with finite mean.

- If $1 < \alpha < 2$, the two-hop relaying algorithm, introduced by [11], is not stable in the sense that the delay incurred has an infinite expectation. It is however still possible to build a naive forwarding algorithm that achieves a delay with finite mean. This requires that a number of m duplicate copies of the data are produced and forwarded, where m must be greater than $\frac{1}{\alpha-1}$, and the network must contain at least $\frac{2}{\alpha-1}$ devices.
- If $\alpha < 1$, none of these algorithms, including flooding, can achieve a transmission delay with a finite expectation.

IV. RELATED WORK

Our opportunistic communication model is related to both Delay-Tolerant Networking and Mobile Ad-Hoc Networking[3]. Research works on MANET and DTN confirm the importance of the problem we address, as several proposition were made to use mobile devices as relays for data transport. Such an approach was considered to enable communication where no contemporaneous path may be found [8], to gather efficiently information in a network of low power sensors [15], [16], [17], or to improve the spatial reuse of dense MANET [11], [12], [18], [13]. All these works prove that the mobility model that is assumed has a strong impact on the performance of the algorithms proposed.

We did not find any previous work studying the characteristics of inter-contact time for users of portable wireless devices. However, we have identified related work in the area of modeling and forwarding algorithms.

A common property of many mobility models found in the literature is that the tail of the inter-contact distribution decays exponentially. In other words, for these models, the inter-contact time is light tailed. This is the case for i.i.d. location of devices in a bounded region (as assumed in [11]), or in the case of the popular random way-point model as demonstrated in [18]. It was shown in a recent article [13] that, by opposition, devices moving according to a Brownian motion, in a bounded region, exhibit heavy tailed inter contact time, with a finite variance (corresponding in our analysis to the case $\alpha > 2$).

The most relevant work is the algorithm proposed by Grossglauser and Tse in [11], further analyzed in [12][18][13]. The two-hop relay forwarding algorithm was initially introduced to study how the mobility of devices impacts the capacity of the network. Our work starts from very different assumptions, as bandwidth

[3]www.dtnrg.org and www.ietf.org/html.charters/manet-charter.html

might be unlimited at each contact, and the focus of the analysis is on the delay incurred by the data transported.

V. SUMMARY, CONCLUSION AND FUTURE WORK

We study a scenario where mobility of networked devices and the opportunistic connection with other devices are used to transfer data. We observe from six experimental traces that the distribution of the inter-contact time seen between two devices in an opportunistic networking environment exhibits a *heavy tail* over a large range of value, that can be compared to a power law with a coefficient less than one. This observation is in contrast with the exponential decay assumption made implicitly by mobility models used to date in ad-hoc networking.

We prove the following major result: Naive forwarding algorithm may deliver data with a bounded expected delay in the case of light tailed inter-contact times, as well as when mobility of devices implies power law inter-contact with coefficient greater than 1. But all of these algorithms have indeed an infinite expected delay when mobility implies power law with coefficient smaller than 1.

Some of the implications of our findings are:

1) Current mobility models (e.g. random way-point, uniformly distributed locations) do not exhibit characteristics found in our six data sets. New mobility models are therefore required.
2) Little work has been done in the area of informed design of opportunistic forwarding algorithms — this remains an area for study. Suitable directions for work might involve the sharing of recent contact information between devices, leading to a more careful selection of potential relay devices which are likely to have a short path to the destination, while also being independently moving as compared to other chosen relay devices.

We will investigate these various aspects of forwarding in opportunistic networks.

We also intend to perform more human mobility experiments. The power law nature of the inter-contact time distribution seems well established for mutual sightings of devices carried by humans during the working day, or in a conference. But parameters in the environment or in the nature of transfer opportunities (including special devices, or infrastructure) certainly affect the shape of the distribution. These new experimental data will allow us to revisit three assumptions made in this paper:

First, approximating the inter-contact time by a power law does not seem to be valid over an unlimited range. Experimental results indicate that the inter-contact time

distribution may be different for very large time scale. However, given that the tail is impacted by the experiment duration, it is difficult to say if the tail of the distribution is a characteristic of the user mobility or a side-effect of the experimental methodology. If this different behavior occurring at very large time scale is confirmed, it may avoid the unbounded expectation of the delay observed in our model. One should nevertheless keep in mind that the range of values where the nature of the distribution may change may be far above the delays that most network applications can tolerate.

Our model also assumes that the contact process of a pair of devices follows a renewal process with a given inter-contact time distribution. It may be possible to get rid of this assumption, assuming that the inter contact times follow locally in time a stationary process that may exhibit some memory in the sequence distribution. Another exciting direction is to approximate the inter-contact times with phase-type distribution. We have indeed already observed that the parameter of the distribution of the inter-contact may change with the time of the day [7].

Similarly, we assumed that contact processes for different pairs of devices are independent, and all described by the same distribution. This is certainly a major simplification. These assumptions are usual, either explicitly or implicitly, in the modeling literature of mobile ad-hoc networks. However, they might not be very realistic and this remains a large area of investigation for opportunistic communication. The contact process of two people is certainly different depending on the communities that they have in common explicitly (work group, institution, friends), or implicitly (nearby neighbors).

VI. ACKNOWLEDGEMENTS

REFERENCES

[1] M. McNett and G. M. Voelker, "Access and mobility of wireless PDA users," Computer Science and Engineering, UC San Diego, Tech. Rep., 2004.

[2] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *MobiCom '04: Proceedings of the 10th annual international conference on Mobile computing and networking*, 2004, pp. 187–201.

[3] J. Su, A. Chin, A. Popivanova, A. Goel, and E. de Lara, "User mobility for opportunistic ad-hoc networking," in *Proceedings of the 6th IEEE Workshop on Mobile Computing Systems and Applications (WMCSA'04)*, 2004.

[4] P. Hui, A. Chaintreau, J. Scott, R. Gass, J. Crowcroft, and C. Diot, "Pocket switched networks and the consequences of human mobility in conference environments," in *Proceedings of ACM SIGCOMM first workshop on delay tolerant networking and related topics (WDTN-05)*, Aug 2005.

[5] P. Bremaud, *Markov Chains, Gibbs Field, Monte Carlo Simulation and Queues*. Springer-Verlag, 1999.

[6] F. Baccelli and P. Bremaud, *Elements of Queuing Theory*, 2nd ed. Springer-Verlag, 2003.

[7] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Pocket switched networks: Real-world mobility and its consequences for opportunistic forwarding," University of Cambridge, Computer Lab, Tech. Rep. UCAM-CL-TR-617, Feb 2005. [Online]. Available: http://www.cl.cam.ac.uk/TechReports/UCAM-CL-TR-617.html

[8] A. Vahdat and D. Becker, "Epidemic routing for partially-connected ad hoc networks," University of California San Diego, Tech. Rep. CS-2000-06, Jul 2000.

[9] X. Chen and A. Murphy, "Enabling disconnected transitive communication in mobile ad hoc networks," in *Proceedings of the Workshop on Principles of Mobile Computing*, Aug 2001, pp. 21–27.

[10] J. A. Davis, A. H. Fagg, and B. N. Levine, "Wearable computers as packet transport mechanisms in highly-partitioned ad-hoc networks," in *ISWC '01: Proceedings of the 5th IEEE International Symposium on Wearable Computers*, 2001, p. 141.

[11] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Transactions on Networking*, vol. 10, no. 4, pp. 477–486, 2002.

[12] G. Sharma and R. Mazumdar, "Scaling laws for capacity and delay in wireless ad hoc networks with random mobility," in *IEEE International Conference on Communications*, vol. 7, Jun 2004, pp. 3869–3873.

[13] X. Lin, G. Sharma, R. R. Mazumdar, and N. B. Shroff, "Degenerate delay-capacity trade-offs in ad hoc networks with brownian mobility," mar 2005, preprint, School of ECE, Purdue Univeristy. [Online]. Available: http://min.ecn.purdue.edu/ linx/paper/it05.pdf

[14] J. F. C. Kingman, "Inequalities in the theory of queues," *J. Roy. Statistical Society, Series B*, vol. 32, pp. 102–110, 1970.

[15] S. Jain, R. C. Shah, G. Borriello, W. Brunette, and S. Roy, "Exploiting mobility for energy efficient data collection in sensor networks," in *WiOpt*, March 2004.

[16] P. Juang, H. Oki, Y. Wang, M. Martonosi, L. S. Peh, and D. Rubenstein, "Energy-efficient computing for wildlife tracking: design tradeoffs and early experiences with zebranet," in *ASPLOS-X: Proceedings of the 10th international conference on Architectural support for programming languages and operating systems*, 2002, pp. 96–107.

[17] T. Small and Z. J. Haas, "The shared wireless infostation model: a new ad hoc networking paradigm (or where there is a whale, there is a way)," in *MobiHoc '03: Proceedings of the 4th ACM international symposium on Mobile ad hoc networking & computing*, 2003, pp. 233–244.

[18] G. Sharma and R. R. Mazumdar, "Delay and capacity trade-off in wireless ad hoc networks with random way-point mobility," mar 2005, preprint, School of ECE, Purdue Univeristy. [Online]. Available: http://min.ecn.purdue.edu/ linx/paper/it05.pdf

A. Preliminary Results

1) Independent composition and limit expectation:
The following lemma will be useful to prove limit result on compound process:

Lemma 2 *Let* $((F_t^{(i)})_{t\in\mathbb{N}})_{i\in I}$ *be a finite collection of sequences of real valued random variables verifying,* $\lim_{t\to\infty}\mathbb{E}\left[F_t^{(i)}\right] = l$, *where* $l \in \mathbb{R}\cup\{+\infty\}$, *and*

(a) $\forall i,t,$ $\mathbb{E}\left[F_t^{(i)}\right] \in \mathbb{R}$, *and* $l \in \mathbb{R}$,

or (b) $\forall i,t,$ $\mathbb{E}\left[F_t^{(i)}\right] \in \mathbb{R}\cup\{+\infty\}$ *and* $l = +\infty$.

Let $(t_k)_{k\in\mathbb{N}}$ *and* $(i_k)_{k\in\mathbb{N}}$ *be two* $\mathbb{N}$ *valued processes, independent from* F, *such that* $\lim_{k\to\infty}t_k = +\infty$ *a.s.* .

We then have $\lim_{k\to\infty}\mathbb{E}\left[F_{t_k}^{(i_k)}\right] = l$

Proof: Let us first develop the following expectation

$$\mathbb{E}\left[F_{t_k}^{(i_k)}\right] = \sum_{i\in I}\sum_{t\geq 0}\sum_{j\geq 0}j\mathbb{P}\left[i_k = i, t_k = t, F_t^{(i)} = j\right]$$

$$= \sum_{i\in I}\sum_{t\geq 0}\sum_{j\geq 0}j\mathbb{P}\left[i_k = i\right]\mathbb{P}\left[t_k = t\right]\mathbb{P}\left[F_t^{(i)} = j\right]$$

$$= \sum_{i\in I}\sum_{t\geq 0}\mathbb{P}\left[i_k = i\right]\mathbb{P}\left[t_k = t\right]\mathbb{E}\left[F_t^{(i)}\right] \tag{3}$$

If we suppose (a), we have $l < +\infty$ and

$$\forall\varepsilon > 0, \exists T \text{ s.t. } (t > T \implies \left|\mathbb{E}\left[F_t^{(i)}\right] - l\right| < \frac{\varepsilon}{2}) .$$

Let $M = \sup_{i\in I, t\leq T}\left|\mathbb{E}\left[F_t^{(i)}\right] - l\right|$, there exists K s.t.

$$k > K \implies \mathbb{P}\left[t_k \geq T\right] \geq 1 - \frac{\varepsilon}{2.M} \text{ and hence}$$

$\left|\mathbb{E}\left[F_{t_k}^{(i_k)}\right] - l\right|$ can be bounded from above by

$$\sum_{i\in I}\sum_{t\geq 0}\mathbb{P}\left[i_k = i\right]\mathbb{P}\left[t_k = t\right]\left|\mathbb{E}\left[F_t^{(i)}\right] - l\right|$$

$$\leq \sum_{i\in I}\mathbb{P}\left[i_k = i\right]\ \left(M.\sum_{t\leq T}\mathbb{P}\left[t_k = t\right]\right.$$

$$\left. + \sum_{t > T}\mathbb{P}\left[t_k = t\right]\left|\mathbb{E}\left[F_t^{(i)}\right] - l\right|\right)$$

$$\leq \sum_{i\in I}\mathbb{P}\left[i_k = i\right](\varepsilon/2 + \varepsilon/2) \leq \varepsilon .$$

Let us now suppose (b), we have $l = +\infty$ and

$$\forall A > 0, \exists T, (t > T \implies \mathbb{E}\left[F_t^{(i)}\right] \geq 2.(A+1)) .$$

Let $M' = \sup_{i\in I, t\leq T}\max\left(-\mathbb{E}\left[F_t^{(i)}\right], 0\right)$. $\exists K$ s.t.:

$$k > K \implies \mathbb{P}\left[t_k \geq T\right] \geq \max\left(1/2, 1 - 1/M'\right), \text{ and}$$

$$\mathbb{E}\left[F_{t_k}^{(i_k)}\right] = \sum_{i\in I}\sum_{t\geq 0}\mathbb{P}\left[i_k = i\right]\mathbb{P}\left[t_k = t\right]\mathbb{E}\left[F_t^{(i)}\right]$$

$$\geq \sum_{i\in I}\mathbb{P}\left[i_k = i\right]\ \left(-M.\sum_{t\leq T}\mathbb{P}\left[t_k = t\right]\right.$$

$$\left. + \sum_{t > T}\mathbb{P}\left[t_k = t\right]\mathbb{E}\left[F_t^{(i)}\right]\right)$$

$$\geq \sum_{i\in I}\mathbb{P}\left[i_k = i\right]\left(-1 + \frac{1}{2}(2(A+1))\right) \geq A .$$

∎

2) Remaining inter-contact: Because the contact process $(U_t^{(d,d')})_{t\geq 0}$ is a renewal process, the sequence $(R_t^{(d,d')})_{t\geq 0}$ of integers is an Homogeneous Markov Chain in $\mathbb{N}$ such that:

$$\begin{cases} R_{t+1}^{(d,d')} = R_t^{(d,d')} - 1 & \text{if } R_t^{(d,d')} > 0, \\ R_{t+1}^{(d,d')} = i - 1 \text{ with prob. } \mathbb{P}\left[X = i\right] & \text{if } R_t^{(d,d')} = 0. \end{cases} \tag{4}$$

This Markov Chain is clearly irreducible and aperiodic as $\mathbb{P}\left[X = 1\right] > 0$, it is recurrent as X is almost surely finite. The following lemma characterizes its properties, which depend on the value of α, based on classical results from the theory of Markov chains.

Lemma 3 *For any devices* d, d', e, e' *we have*

(i) *If* $\alpha > 1$, $(R_t^{(d,d')})_{t\geq 0}$ *is recurrent positive.*

(ii) *If* $\alpha > 1$, *the chain* $(R_t^{(d,d')}, R_t^{(e,e')})_{t\geq 0}$ *is ergodic and admits the following stationary distribution:*

$$\pi(i,j) = \frac{(i+1)^{-\alpha}(j+1)^{-\alpha}}{(c_1)^2} \text{ where } c_1 = \sum_{i'\geq 0}(i'+1)^{-\alpha} .$$

(iii) *If* $\alpha \leq 1$, $(R_t^{(d,d')})_{t\geq 0}$ *is recurrent null.*

Proof: Let us introduce $\mathtt{ret}_0$ the time for $R^{(d,d')}$ to return in the state 0. From the structure of the Markov chain (4), starting from state 0, we can easily deduce that $\mathbb{E}_0[\mathtt{ret}_0] = \mathbb{E}[X]$. If $\alpha > 1$, we have $\mathbb{E}[X] < +\infty$, proving (i), and if $\alpha \leq 1$, we have $\mathbb{E}[X] = +\infty$, proving (iii).

By (i), we know that the Markov chain $R^{(d,d')}$ is recurrent positive, hence it admits a stationary distribution. It is easy to check, from its regenerative structure, that it is given by: $\pi(i) = c_1(i+1)^{-\alpha}$ where $c_1 = 1/\sum_{i\geq 0}(i+1)^{-\alpha}$.

The same result holds for $R^{(e,e')}$. As these two Markov Chains are independent, one can then check easily that the product Markov chain $(R^{(d,d')}, R^{(e,e')})$, which is irreducible and aperiodic, admits a stationary distribution given by the product of the measure. It is hence ergodic.

∎

Smith's formula for $\alpha > 1$**:** For any devices d and d', the process $(R_t^{(d,d')})_{t\geq 0}$ is regenerative with respect to the delayed renewal sequence $(T_k^{(d,d')})_{k\geq 0}$. If we assume

$\alpha > 1$, we have $\mathbb{E}[X] < +\infty$, hence the inter-event of the sequence $(T_k^{(d,d')})_{k \geq 0}$ admits a finite mean. We know in this case (see [5] p.148) that

$$\lim_{t \to \infty} \mathbb{E}\left[f(R_t^{(d,d')}) \right] = \frac{\mathbb{E}\left[\sum_{t=T_0^{(d,d')}}^{T_1^{(d,d')}-1} f(R_t^{(d,d')}) \right]}{\mathbb{E}\left[T_1^{(d,d')} - T_0^{(d,d')} \right]}$$

for any f verifying $\mathbb{E}\left[\sum_{t=T_0^{(d,d')}}^{T_1^{(d,d')}-1} |f(R_t^{(d,d')})| \right] < \infty$. (5)

B. Proof of Lemma 1

Lemma 1 is a generalization of the method presented in § III-C. Let us start by the following remark:

for $\beta < \alpha$, $\sum_{i=1}^{j} i^{\beta-\alpha} \leq \frac{(\beta-\alpha)+j^{\beta-\alpha+1}}{\beta-\alpha+1}$. (6)

Indeed, the function $x \mapsto x^{\beta-\alpha}$ is non-increasing on $]0; +\infty[$, hence we have for any $i \geq 2$,

$$i^{\beta-\alpha} \leq \int_{i-1}^{i} x^{\beta-\alpha} dx < +\infty \text{ hence (6) follows from}$$

$$\sum_{i=1}^{j} i^{\beta-\alpha} \leq 1 + \int_1^j x^{1-\alpha} dx = 1 + \frac{j^{\beta-\alpha+1}-1}{\beta-\alpha+1}.$$

As all processes of contacts between devices are independent, the stationary distribution of the product of m Markov Chains is given by the the product measure. Hence we have that $\mathbb{E}\left[\min\left(R_t^{(d_1,d'_1)}, \ldots, R_t^{(d_m,d'_m)} \right) \right]$ is equal to $\frac{1}{(c_1)^m} \sum_{i_1,\ldots,i_m} \frac{\min(i_1,\ldots,i_m)}{(i_1+1)^\alpha \ldots (i_m+1)^\alpha}$. In particular this value is finite if we can prove $g(m,\alpha,1) < \infty$, where g is defined as

$$g(m,\alpha,\beta) = \sum_{i_1,\ldots,i_m} \frac{(\min(i_1,\ldots,i_m))^\beta}{(i_1)^\alpha \ldots (i_m)^\alpha}.$$

We will prove more generally that if $\alpha > 1 + 1/m$, and $\beta \leq 1$, then $g(m,\alpha,\beta) < \infty$.

For $m = 1$, this is true as for $\alpha > 2$ and $\beta \leq 1$, $g(1,\alpha,\beta) = \sum_i i^{\beta-\alpha} < +\infty$.

More generally, $g(m,\alpha,\beta)$ is bounded by

$$m. \sum_{i_2,\ldots,i_m} \left(\sum_{i_1=1}^{\min(i_2,\ldots,i_m)} (i_1)^{\beta-\alpha} \right) \frac{1}{(i_2)^\alpha \ldots (i_m)^\alpha}$$

$$\leq m. \sum_{i_2,\ldots,i_m} \frac{a_m + b_m(\min(i_2,\ldots,i_m))^{\beta+1-\alpha}}{(i_2)^\alpha \ldots (i_m)^\alpha}$$

$$\leq a'_m + b'_m g(m-1,\alpha,\beta+1-\alpha)$$

$$\leq \ldots$$

$$\leq a' + b' \underbrace{g(1,\alpha,\beta+(m-1)(1-\alpha))}_{=\sum_i i^{\beta+(m-1)(1-\alpha)-\alpha}}.$$

a_m, b_m, a'_m, b'_m are finite constant real number, that depends on α, β, and m. They can be computed using (6); a and b can also be computed, by developing the recurrence equation, but the exact values of these constants have no importance for the result of the theorem. The result follows as

$$\beta + (m-1)(1-\alpha) - \alpha \leq \underbrace{\beta-2}_{<-1} + \underbrace{m+1-m.\alpha}_{<0}.$$

$\blacksquare$

For any real numbers $(x_1,\ldots,x_m)$, and $i \leq m$, let us denote by $\texttt{ord}(i,(x_1,\ldots,x_m))$ the i_th element of the sequence after it is reordered in the increasing order. In particular $\texttt{ord}(i,(x_1,\ldots,x_m)) = \min(x_1,\ldots,x_m)$. We have

Corollary 1 *Let* $(R_t^{(d_1,d'_1)})_{t \geq 0}, \ldots, (R_t^{(d_m,d'_m)})_{t \geq 0}$ *be the remaining inter-contact times for m different pairs of devices* $(d_i,d'_i)_{1 \leq i \leq m}$. *We suppose that* $\alpha > \frac{(m-j+1)+1}{m-j+1}$,

then $\mathbb{E}\left[\texttt{ord}\left(j, (R_t^{(d_1,d'_1)}, \ldots, R_t^{(d_m,d'_m)}) \right) \right] < \infty.$

Proof: Following the same step than for the previous proof we can easily show that this expectation is given by :

$$\frac{1}{(c_1)^m} \sum_{i_1,\ldots,i_m} \frac{\texttt{ord}(j,(i_1,\ldots,i_m))}{(i_1+1)^\alpha \ldots (i_m+1)^\alpha}.$$

By symmetry we can show that this sum is upper bounded by

$$\frac{1}{(c_1)^m} m \sum_{i_1,\ldots,i_m} \frac{i_j \mathbb{I}_{\{\texttt{ord}(j,(i_1,\ldots,i_m))=i_j\}}}{(i_1+1)^\alpha \ldots (i_m+1)^\alpha}.$$

Considering the number of choices of $j-1$ elements in $m-1$, we can show by another symmetry argument,

$$\leq c_2 \sum_{i_1,\ldots,i_m} \frac{\texttt{ord}(j,(i_1,\ldots,i_m))}{(i_1+1)^\alpha \ldots (i_m+1)^\alpha}$$
$$\times \mathbb{I}_{\{\texttt{ord}(j,(i_1,\ldots,i_m))=i_j, i_j \leq i_{j+1},\ldots,i_j \leq i_m\}}.$$

with $c_2 = m \frac{1}{(c_1)^m} \frac{(m-1)!}{(j-1)!(m-1)!}$. In the term of this sum, i_j is always the maximum of $i_1,\ldots,i_{j-1}, i_j$ and the minimum of $i_j, i_{j+1}, \ldots, i_m$. The product of terms corresponding to $i_1,\ldots,i_{j-1}$ is taken only for them with values in $\{1,2,\ldots,i_j\}$; it could be overall upper bounded by c_1^{j-1}, by completing each sum. We can then show that

$$\leq c_2 c_1^{j-1} \underbrace{\sum_{i_j,\ldots,i_m} \frac{\min(i_j,\ldots,i_m)}{(i_j)^\alpha \ldots (i_m)^\alpha}}_{g(m-j+1,\alpha,1)}.$$

The result is then proved, once we remember, from the proof of Lemma 1, that $g(m-j+1,\alpha,1) < +\infty$ if $\alpha \geq 1 + \frac{1}{m-j+1}$.

$\blacksquare$