Consiglio Nazionale delle Ricerche

Analysis of individual pair and aggregate inter-contact times in heterogeneous opportunistic networks

A. Passarella, M. Conti

IIT TR-18/2011

Technical report

settembre 2011

Istituto di Informatica e Telematica

Analysis of individual pair and aggregate inter-contact times in heterogeneous opportunistic networks

Andrea Passarella and Marco Conti, *Member, IEEE*

Abstract—Foundational work in the area of opportunistic networks has shown that the distribution of inter-contact times between pairs of nodes has a key impact on the network properties, e.g. in terms of convergence of forwarding protocols. Specifically, forwarding protocols may yield infinite expected delay if the inter-contact time distributions present a particularly heavy tail. While these results hold for the distributions of inter-contact times between *individual pairs*, most of the literature uses the *aggregate* distribution, i.e. the distribution obtained by considering the samples from all pairs together, to characterise the properties of opportunistic networks. In this paper we analyse when this approach is correct and when it is not. We study, through an analytical model, the dependence between the individual pair and the aggregate distributions. We show that the aggregate distribution can be way different from the distributions of individual pair inter-contact times. Therefore, using the former to characterise properties that depend on the latter is not correct in general, although this is correct in some cases. We substantiate this finding by analysing the most representative distributions characterising real opportunistic networks that have been reported in the literature based on trace analysis. We study networks whose aggregate inter-contact time distribution presents a heavy tail with or without exponential cutoff. We show that a exponential cutoff in the aggregate appears when the average inter-contact times of individual pairs are finite. We also show that, when individual pairs follow Pareto distributions, the aggregate distribution consistently presents a heavy tail. However, heavy tail aggregate distributions can also emerge in networks where individual pair inter-contact times are not heavy tailed, e.g. exponential or Pareto with exponential cutoff distributions. This constitutes a reassuring result, as it means that forwarding protocols do not necessarily diverge in the quite common case of networks whose aggregate inter-contact time distribution is heavy tailed.

Index Terms—opportunistic networks, analytical modelling

1 Introduction

Opportunistic networks [1] are mobile self-organising networks where the existence of a continuous multi-hop path formed by simultaneously connected hops is not taken for granted. In order to deliver a message from a source to a destination, in opportunistic networks it is required that a *space-time* multi-hop path exists [2], [3] (see Figure 1 for a graphical example). Due to users' mobility and network reconfigurations, different portions of a space-time path can become available at different points in time. For example, in Figure 1 node 2 moves close to node 3 at time t_2, while node 5 moves close to the destination at time t_3, thus establishing a space-time path between nodes S and D. Intermediate nodes in space-time paths exploit the store-carry-and-forward concept [4], [5]: They temporarily store messages addressed to a currently unreachable destination (if "better" next hops are currently not available), until a new portion of the space-time path appears, and therefore the message can progress towards the final destination.

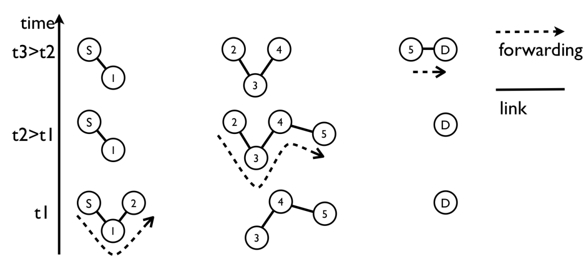

Fig. 1. Example of a space-time path

Foundational results in the area of opportunistic networks have clearly shown that characterising inter-contact times between nodes is crucial [6], [7], [8]. Starting from the point in time when two nodes loose single-hop connectivity (i.e., a contact finishes), an inter-contact time is the time until they are able to directly communicate again (i.e. a contact starts). As in opportunistic networks contacts are the only way for messages to progress towards the destination, the distribution of inter-contact times plays a key role in determining the performance of forwarding

• A. Passarella and M. Conti are with IIT-CNR (Institute for Informatics and Telematics of the National Research Council), Via G. Moruzzi 1, 56124 Pisa, Italy
E-mail: firstname.lastname@iit.cnr.it

This work was partially funded by the European Commission under the FIRE SCAMPI (FP7-IST-258414), and FET-AWARE RECOGNITION (FP7-IST-257756) Projects.

protocols[1]. Specifically, Chaintreau et al. [6] show that in a homogeneous network where inter-contact times between all pairs of nodes are iid (independent identically distributed), if inter-contact times are heavy tailed with coefficient $\alpha < 2$ an important class of forwarding protocols (termed "naïve") diverge, i.e. yield infinite expected delay. In naïve protocols, nodes do not exploit any information describing the status of the network when taking forwarding decisions, and are only aware of some identifier of the destination, such as its address. These protocols are attractive because they are very lightweight and simple to implement and analyse, and have been widely used in the literature [9], [10], [11], [12], starting from the seminal work on Epidemic routing [13]. Notably, the 2-hop forwarding protocol used in [10] to derive foundational results on the capacity of opportunistic networks falls in this category. Doubts have been raised [7], [8] about the fact that inter-contact times in popular traces [14], [15], [16], [17] actually meet the conditions found in [6] for divergence of protocols (see Section 2). While Chaintreau et al. [6] argue that the inter-contact times of these traces can be well approximated with a Pareto distribution (thus meaning they are power law), [7], [8] note that a Pareto distribution with a final exponential cutoff may be a better approximation (meaning that the inter-contact times are not power law). However, this debate does not challenge the theoretical result proven in [6] about the fact that naïve forwarding protocols may diverge depending on the inter-contact time distributions. Based on these results, it is clear that Pareto distributions with or without exponential cutoff are particularly important for the analysis of inter-contact times. We will also use them extensively in the following of the paper. Specifically, we will consider Pareto distributions whose Complementary Cumulative Distribution Function (CCDF, i.e. $P(X > x)$) is in one of the following forms:

$$P(X > x) = \left(\frac{b}{b+x}\right)^{\alpha}, \ \alpha, b, x > 0 \tag{1}$$

$$P(X > x) = \left(\frac{b}{x}\right)^{\alpha}, \ \alpha, b > 0, x > b \tag{2}$$

where α is the "shape" parameter and b the "scale" parameter. The form in Equation 1 allows values arbitrarily close to 0, while the form in Equation 2 does not. This has important implications, as we discuss in the paper. In the following, we will refer to the former as "Pareto0" and to the latter as "Pareto". Finally, we will also consider Pareto distributions with exponential cutoff, whose CCDF is in the form [18]

$$P(X > x) = \frac{\Gamma(1-\alpha, \mu x)}{\Gamma(1-\alpha, \mu b)}, \ \alpha > 1, \ \mu, b, x > 0 \ , \tag{3}$$

where $\Gamma(s, x) = \int_x^\infty t^{s-1} e^{-s} dt$ is the upper incomplete Gamma function [19], α, b and μ are the shape, scale, and rate parameters, respectively.

There is significant ambiguity in the literature on whether the distributions of inter-contact times of *individual pairs* or the distribution of the *aggregate* inter-contact times should be used to characterise opportunistic networks, the aggregate distribution being the distribution of inter-contact times of *all* pairs considered together. Typically in the literature [6], [8], [20], [21], [22], [23] the aggregate distribution is used. From a practical standpoint the aggregate distribution is more manageable, as less samples are required to characterise its statistics with respect to those of all individual pair distributions, and only one distribution can be used to characterise the entire network. Being able to use the aggregate distribution instead of all distributions of individual pairs would be, therefore, desirable. However, this is not always possible. Although the results obtained by Chaintreau et al. [6] hold for the distributions of the individual pairs, authors correctly use the aggregate distribution. This is because they assume a homogeneous network[2], i.e. one where all individual pair distributions are iid, and thus the aggregate distribution is the same as the (common) distribution of individual pairs. However, as we show in this paper, this is not correct in general if the network is *heterogeneous*. This remark has been significantly overlooked in the opportunistic networking literature, which mostly assumes that i) the distribution of aggregate inter-contact times well represents the distributions of all individual pair inter-contact times, ii) all these distributions are power law (based on the results in [6]) or power law with an exponential cutoff (based on the results in [8], [7]), and iii) in a network with a power law aggregate inter-contact time distribution, naïve forwarding protocols diverge. For example, the validation of several reference mobility models [20], [21], [22], [23] is carried out by showing that either the distributions of individual pair inter-contact times or the aggregate distribution follow a power law with or without an exponential cutoff. These beliefs are very well established. There are a few (to the best of our knowledge) papers [24], [25] that analyse the distributions of individual pairs in reference traces finding a good fit with *exponential* distributions for a significant fraction of them. These papers had a rather limited impact, as they seem to contradict the hypothesis of inter-contact times being power law.

1. In opportunistic networks the routing and forwarding process are carried out at the same time and are implemented by a unique algorithm. Therefore in the following we use the terms routing and forwarding interchangeably.

2. In principle, it would be more precise to use the term "network graph" instead of "network" in this case. Hereafter, we use the two terms interchangeably, when the meaning is clear from the context.

In this paper we clarify the dependence between the individual pair inter-contact time distributions and the aggregate distribution through an analytical model. The model allows us to challenge these common beliefs, and shows that overlooking the difference between the individual pair and the aggregate distributions can lead to completely wrong conclusions. We are able to show, among others, that in several relevant cases naïve forwarding protocols will *not* diverge even though the aggregate distribution is power law.

We consider a heterogeneous networking environment, in which the individual pair distributions are of the same type (e.g., exponential, Pareto, ...), but whose parameters can be different from one pair to another, and are unknown a-priori. In this case, individual pair inter-contact times are not identically distributed. We assume that the contact *rate* between a pair (the reciprocal of the pair average inter-contact time) is drawn from a given distribution, which, therefore, determines the specific parameters of the pair inter-contact time distribution. In other words, as the distribution of the rates controls the parameters of the inter-contact time distributions, it allows us to control the heterogeneity of the network.

The model described in the paper shows that both the distribution of the rates and the distributions of individual pair inter-contact times impact on the aggregate distribution. In particular, we use the model to find, among others, the conditions under which the aggregate distribution features the main characteristics often found in traces, i.e. a power-law with or without exponential cutoff. We can summarise the key findings presented in the paper as follows.

- Starting from *exponentially* distributed individual pair inter-contact times, the aggregate is distributed *exactly* according to a Pareto law iff the contact rates are drawn from a Gamma distribution.
- When individual pair inter-contact times are *exponential*, and rates are drawn from a Pareto distribution, the *asymptotic* behaviour of the aggregate distribution (for large inter-contact times) is a power-law *with or without* exponential cutoff. In particular, the long tail behaviour appears when rates *can* be arbitrarily close to 0, i.e. when average inter-contact times can tend to infinity.
- When pair inter-contact times follow a Pareto distribution with fixed shape or scale parameters, the aggregate distribution is power law for a significant range of contact rate distributions. In particular, when the shape parameter is fixed, the aggregate is power law with the same exponent no matter what distribution of rates.
- When pair inter-contact times follow a Pareto distribution with exponential cutoff, the aggregate distribution can present exactly the same shape under certain conditions. It is however power law *without* exponential cutoff, for contact rate distributions allowing rates arbitrarily close to 0.

The contribution of this paper is thus manifold. Besides providing a detailed model that describes the dependence between individual pair and aggregate inter-contact time distributions, our findings reconcile apparently contradicting results previously found in the literature by analysing real traces. Specifically, our results show that exponentially distributed individual pairs (found in [25], [24]) are compatible with power law aggregate inter-contact times with or without an exponential cutoff (found in [6], [8], [7] and then assumed in most of the literature). Moreover, they also show that relying only on the aggregate inter-contact time distribution for assessing key properties of opportunistic networks is not correct in general, and may lead to wrong conclusions. In particular, finding a power-law in the aggregate inter-contact time distribution is not necessarily an indication that individual pair distributions feature a heavy tail as well, and that therefore naïve forwarding protocols may diverge. On the contrary, the heterogeneity of the network, represented in our study by the distribution of the contact rates, plays a crucial role in determining the nature of the aggregate distribution, which can be totally different from the distributions of the individual pairs.

The rest of the paper is organised as follows. We review the relevant state-of-the-art in Section 2. Then, Section 3 presents the general model describing the dependence between the individual pair inter-contact times, the distribution of contact rates, and the aggregate inter-contact time distribution. In Sections 5 and 6 we focus on some of the most relevant cases of inter-contact time distributions found in real traces, and study which types of aggregate distributions emerge depending on the individual pair distributions and the heterogeneity of the network. Together with analytical results, we also present simulation results validating the analytical findings. Finally, in Section 7 we draw the main conclusions of this study.

2 RELATED WORK

The first body of work, to the best of our knowledge, that highlights the importance of inter-contact times for characterising opportunistic networks was carried out in the framework of the EU Haggle project [26], [6]. As discussed in Section 1, Chaintreau et al. [6] find very important theoretical results showing that naïve forwarding protocols may diverge in homogeneous networks if individual pair inter-contact times are heavy tailed. Actually they also analyse a popular set of traces [14], [15], [16], [17] finding that the aggregate distribution can be approximated with a Pareto distribution with shape less than one. The straightforward conclusion is that naïve forwarding protocols can easily diverge in real opportunistic networks.

This very pessimistic result is somewhat softened by Karagiannis et al. [7], [8], who re-analyse the same traces and note that the aggregate inter-contact time distribution might indeed present an exponential cutoff in the tail, following the Pareto shape highlighted in [6]. Assuming, again, that the analysed networks are homogeneous, they conclude that naïve forwarding protocols might actually not yield infinite delay. In this work authors discuss the fact that the aggregate and the individual pair distributions may be different. They propose an initial model for studying the dependence between the two, which we exploit as a starting point in our paper. However, they do not study this aspect further, after checking that, in the analysed traces, a subset of individual pair inter-contact times follow a power-law distribution with exponential cutoff.

The above papers informed most of the subsequent literature, which most of the time assumes that the distributions of individual pairs and the aggregate distribution can be used interchangeably. Only a few papers pay attention to individual pair distributions. Among them, Conan et al. [24] re-analyse again the same set of traces, focusing much more than before on the distributions of individual pair inter-contact times. They clearly show that these networks are actually heterogeneous, and that an exponential distribution fits well a significant fraction of individual pair inter-contact times, while Pareto and Lognormal distributions also show a good fit with other subsets of the pairs. Authors also provide a model similar in spirit to the one we use in our work, to analyse conditions under which exponential individual pair distributions can result in a Pareto aggregate. As we highlight in the following, their model does not incorporate a fundamental aspect, and thus obtains imprecise results. Gao et al. [25] analyse the Reality Mining trace [17], finding that exponential distributions fit over 85% of the individual pair inter-contact times. They do not study the dependence between the individual pairs and the aggregate distribution, though.

With respect to this body of work, in this paper we provide a thorough analysis of the dependence between individual pair and aggregate inter-contact time distributions. The model we derive is more general and accurate than the ones presented in [8], [24]. The model allows us to reconcile apparently contradicting results presented in the literature, such as the fact that individual pair exponential inter-contact times are compatible with power-law aggregate inter-contact times. Moreover, with respect to the existing literature, we exploit the model to re-analyse aggregate inter-contact time distributions found in real traces, i.e. power-law distributions with or without an exponential cutoff. Specifically we are able to show that several combinations of individual inter-contact time and contact rate distributions result in aggregate distributions with these shapes. In several cases, even when the aggregate distribution is power law, the individual distributions are not. Thus, our model allows us to provide reassuring results on the convergence properties of naïve forwarding protocols, because the fact that real networks feature a power-law aggregate inter-contact time distribution does not necessarily mean that they diverge.

This paper extends our previous work in [27]. Specifically, in this paper we analyse a much more extended set of heterogeneous networks exploiting our model, investigating, for example, the dependence between individual pair and aggregate inter-contact time distributions when the former are Pareto with or without an exponential cutoff. These additional results allow us to conclude that power-law aggregate distributions can result both starting from exponentially distributed, or from Pareto distributed (with or without exponential cutoff) individual pair inter-contact times. Moreover, with respect to [27], in addition to the new analytical results, in this paper we present a completely new set of simulation results used to validate the analysis. The simulations presented in [27] have been re-run as described in Section 5.3 to achieve higher statistical confidence.

3 Analytical model of aggregate inter-contact times

In this section we present an analytical model that describes the dependence between the inter-contact times of individual pairs and the resulting distribution of aggregate inter-contact times. This is the starting point for the rest of the analysis.

3.1 Preliminaries

As a first step, it is important to recall a result found by Karagiannis et al. [8], which shows the relationship between the distribution of individual pair inter-contact times and the aggregate distribution, in a network where the parameters of the individual pair distributions are known. Let assume to monitor individual pair inter-contact times for a large time interval T. Let denote with P the number of pairs for which at least one inter-contact time is measured over T. Moreover, denote with $F_p(x)$ the CCDF of inter-contact times of pair $p, p \in \{1, \dots, P\}$, $n_p(T)$ and $N(T)$ being the number of inter-contact times of pair p and the total number of inter-contact times over T, respectively. Finally, denote with θ_p the rate of inter-contact times for pair p (i.e. the reciprocal of the average inter-contact time) and with $\theta = \sum_p \theta_p$ the total rate of inter-contact times. Then, the CCDF of the aggregate inter-contact times $F(x)$ can be expressed as in the following lemma.

Lemma 1: In a network where P pairs of nodes exist for which inter-contact times can be observed, the

CCDF of the aggregate inter-contact times is:

$$F(x) = \lim_{T\to\infty} \sum_{p=1}^{P} \frac{n_p(T)}{N(T)} F_p(x) = \sum_{p=1}^{P} \frac{\theta_p}{\theta} F_p(x) \ . \quad (4)$$

Proof: See [8]. □

Lemma 1 is rather intuitive. The distribution of the aggregate inter-contact times is a mixture of the individual pair distributions. Each individual pair "weights" in the mixture proportionally to the number of inter-contact times that can be observed in any given interval (or, in other words, proportionally to the rate of inter-contact times).

3.2 General results

We now extend the result of Lemma 1 to the case in which the parameters of the individual pair inter-contact times are *not* known a priori. Specifically, we consider the general case in which the contact rates are iid and distributed according to a continous random variable Λ with density $f(\lambda), \lambda \geq 0$ (for the generic pair p, λ_p denotes its rate). We also assume that all individual pair inter-contact times follow the same type of distribution. For the generic pair p, the distribution parameters are set such that the resulting rate is equal to λ_p. Note that we are able to model heterogeneous networks, as inter-contact time distributions of different pairs are in general different, as their rates are different[3]. With respect to the notation used in Section 3.1, we hereafter denote with $F_\lambda(x)$ the CCDF of the inter-contact times between a pair of nodes whose rate is equal to λ. Under these assumptions, the CCDF of the aggregate inter-contact times becomes as in Theorem 1.

Theorem 1: In a network where the contact rates are distributed with density $f(\lambda)$, the CCDF of the aggregate inter-contact times is as follows:

$$F(x) = \frac{1}{E[\Lambda]} \int_0^\infty \lambda f(\lambda) F_\lambda(x) d\lambda \ . \quad (5)$$

Proof: The complete proof is available in Appendix A, while here we provide an intuitive sketch. As for Equation 4, also Equation 5 can be seen as a mixture of the CCDFs of individual pairs, $F_\lambda(x)$. In this case, however, the rates are unknown a-priori, and are sampled from a r.v. with density $f(\lambda)$. Therefore, all possible components $F_\lambda(x)$ (corresponding to all possible values of the rates, λ) can appear in the mixture. As λ is sampled from a non-negative continuous r.v., the mixture results in an integral over $[0, \infty)$. The term $\lambda f(\lambda) d\lambda$ is the weight in the mixture of component $F_\lambda(x)$. This weight is actually the product of the rate (λ) by the probability of having the component corresponding to that rate in the mixture ($f(\lambda)d\lambda$). This is intuitive, as it means that a particular component weights, in the mixture, proportionally to i) its probability of being in the mixture, and ii) the number of samples it generates in the mixture (i.e., the value of λ). Finally, the denominator $E[\Lambda]$ results from the normalisation of the weights: The total sum of the weights is $\int_0^\infty \lambda f(\lambda) d\lambda$, which is by definition the average value of Λ. □

Note that the aggregate distribution in Equation 5 does not depend on the number of pairs P anymore, unlike the form in Equation 4. This is because under the assumptions of Theorem 1 each pair can be characterised by *any* contact rate λ with a probability $f(\lambda)d\lambda$. As contact rates are distributed according to a continuous random variable, each pair contributes an infinite number of distributions to the aggregate (each one with an infinitesimal weight). Therefore, the aggregate distribution is always made up of an infinite number of components, irrespective of the specific number of pairs in the network. Thus, the model provided by Theorem 1 holds for any number of pairs P.

Generalising Lemma 1 as in Theorem 1 results in a much richer tool for understanding the dependence between individual pair and aggregate inter-contact time distributions. Specifically in the model provided by Theorem 1 the individual pair distributions are not pre-defined, but can be tuned according to the random variable Λ. This allows us to "steer" and control the heterogeneity of the network. As we show in Sections 5 and 6, this model allows us to study the relationship between individual pair and aggregate inter-contact time distributions, by assuming that i) individual pairs are heterogeneous; ii) their inter-contact times follow an arbitrary family of distributions ($F_\lambda(x)$); and iii) their rates follow another arbitrary distribution ($f(\lambda)$). These degrees of flexibility are not provided by the model in Lemma 1.

As a final remark, a similar generalisation was also attempted in [24]. However, the formulation in [24] is not exact, as it does not take into account the fact that, in the mixture defining $F(x)$, distributions of more frequent contact patterns should "weight more" with respect to distributions of less frequent contact patterns. Specifically, in the formulation in [24], the weight associated with each component $F_\lambda(x)$ is the probability of having the component corresponding to rate λ in the mixture, which is not correct. Consider the case of a toy distribution with only two possible rates $\lambda_1 << \lambda_2$, with the same probability. According to the model in [24], the two components will have the same weight in the mixture. However, over any given amount of time, it is clear that the number of observed inter-contact time samples from a pair whose contact rate is λ_2 will be much higher than the number of observed samples from a pair whose contact rate is λ_1. Therefore, the distribution of individual inter-contact

3. Note that, when $F_\lambda(x)$ is defined by more than one parameter, additional conditions besides the rate should be identified to derive all parameters. Our analysis holds true for any definition of such additional conditions, as shown in Section 6.

times corresponding to λ_2 will contribute many more samples to the aggregate, and therefore, intuitively, should weight much more in the mixture than the distribution corresponding to λ_1.

4 Preview of the main results

In Sections 5 and 6 we use the model derived in Theorem 1 to study several networks, where the individual inter-contact times are exponentially distributed (Section 5) or follow a power-law distribution with or without an exponential cutoff (Section 6). Based on the analysis of individual inter-contact times measured in real traces, [8], [25], [24] these are among the most relevant cases to consider. In general, this analysis highlights several aspects. First of all, the model in Theorem 1 allows us to highlight the resulting aggregate inter-contact times distribution that can emerge starting from exponential and power-law (with or without cutoff) distributions, depending on the type of network heterogeneity (i.e. on the distribution of the contact rates). Moreover, according to the results in [6], [8], the cases we consider cover both networks in which naïve forwarding protocols converge (exponential and power-law with exponential cutoff individual pair distributions), and diverge (power-law distributions). The results we present in the following of this paper show in which of these cases the distribution of aggregate inter-contact times can be used as a correct indicator of naïve forwarding protocol divergence, and when it can not.

Specifically, in Section 5 we show that power-law distributions (with or without exponential cutoff) for the aggregate inter-contact times can appear starting from exponentially distributed individual pair inter-contact times. It is possible to obtain even quite heavy tails in the aggregate $\alpha \in (0, 1)$. Using the aggregate in these cases, one would predict divergence of forwarding protocols, which would be totally wrong. The key reason behind this finding is that when the network is heterogeneous, the heterogeneity of the individual pair distributions plays a crucial role in determining the aggregate distribution of the inter-contact times, which may be of a completely different type with respect to the individual pair distributions. In particular, when contact rates can be arbitrarily close to 0, i.e. average inter-contact times can diverge, the aggregate distribution consistently presents a heavy tail, even though the individual pair distributions are not heavy tailed.

We show a similar property in Section 6 when analysing individual pair inter-contact times following a Pareto distribution with an exponential cutoff. Moreover, in Section 6 we show that, under certain conditions, when individual inter-contact times present a power law either with or without exponential cutoff, the aggregate inter-contact time distribution also presents the same shape in the tail. This is an interesting result, as it highlights cases where the aggregate inter-contact times distribution is representative of the distributions of the individual pairs, although the network is heterogeneous.

Summarising, on the one hand, our results show that - unfortunately - in general studying the aggregate distribution is not sufficient. For example - aggregate power laws can appear both in cases where naïve forwarding protocols diverge (individual inter-contact times following a power law distribution) or converge (individual inter-contact times following an exponential or power law with cutoff distribution). On the other hand, our results are reassuring, as they clearly show that an aggregate inter-contact time distribution presenting a power law is not necessarily an indication of naïve forwarding protocols divergence, as commonly assumed in the literature.

5 Networks with exponential individual inter-contact times

In this section we exploit the model provided by Theorem 1 to investigate the dependence between the distributions of individual pair inter-contact times and their aggregate distribution when the former are exponential. Specifically, we assume $F_\lambda(x) = e^{-\lambda x}$, and study how the aggregate CCDF $F(x)$ varies for different distributions of the contact rates, $f(\lambda)$.

The results are hereafter presented as grouped in two classes. Firstly, in Section 5.1, we investigate under which conditions the aggregate inter-contact times follow *exactly* a given distribution. Specifically, we impose that $F(x)$ in Equation 5 is equal to such distribution, and find the corresponding distribution of the contact rates $f(\lambda)$. Then, in Section 5.2 we find additional cases in which it is not possible to exactly map a given aggregate distribution $F(x)$ to a specific rate distribution $f(\lambda)$, but it is possible to identify rate distributions such that the tail of the aggregate follows a certain pattern.

5.1 Exact aggregate inter-contact time distributions

First of all, we wish to identify rate distributions $f(\lambda)$ that result in power-law (Pareto) aggregate distributions. From Equation 5, and recalling that we assume individual inter-contact times are exponentially distributed, we have to find $f(\lambda)$ such that

$$\frac{1}{E[\Lambda]} \int_0^\infty \lambda f(\lambda) e^{-\lambda x} d\lambda = \left(\frac{b}{b+x} \right)^\alpha , \qquad (6)$$

where α and b are the shape and scale parameters of the Pareto distribution. Note that in this case we consider the definition of the Pareto distribution in which all positive values are admitted, i.e., $x > 0$.

The rate distribution $f(\lambda)$ satisfying Equation 6 is provided by Theorem 2. It is worth noting that a qualitatively similar result was also found in [24]. However, due to the inexact formulation of $F(x)$ discussed in Section 3.2, the exact result differs. Specifically, the parameters of the rate distribution found in [24] are different with respect to the ones derived in Theorem 2.

Theorem 2: When individual pair inter-contact times are exponentially distributed, aggregate inter-contact times are distributed according to a Pareto law with parameters $\alpha > 1$ and $b > 0$ iff the contact rates follow a Gamma distribution $\Gamma(\alpha - 1, b)$, i.e.

$$F(x) = \left(\frac{b}{b+x}\right)^{\alpha} \iff f(\lambda) = \frac{b^{\alpha-1}}{\Gamma(\alpha-1)}\lambda^{\alpha-2}e^{-b\lambda} \ . \tag{7}$$

Proof: See Appendix B. □

As discussed in Sections 1 and 2, based on the results in [6] it has been common in the literature to assume that, if the aggregate inter-contact time distribution is Pareto with $\alpha \in (1, 2]$, naïve forwarding protocols yield infinite delay. Theorem 2 clearly shows that this is not correct, as aggregate power-laws with $\alpha \in (1, 2]$ can be obtained starting from exponential individual pair inter-contact times. In such a case, the expected delay of naïve forwarding protocols is finite.

As a special case of Theorem 2, the following corollary holds true.

Corollary 1: When individual pair inter-contact times are exponentially distributed, aggregate inter-contact times are distributed according to a Pareto distribution with parameters $\alpha = 2$ and $b > 0$ iff the rates of individual inter-contact times follow an exponential distribution with rate b, i.e.

$$F(x) = \left(\frac{b}{b+x}\right)^{2} \iff f(\lambda) = be^{-b\lambda} \ . \tag{8}$$

Proof: This follows immediately from Equation 7 by recalling that a Gamma distribution $\Gamma(1, b)$ is actually exponential with rate b. □

Corollary 1 further stresses the result of Theorem 2, stating that a power-law distribution of aggregate inter-contact times can be obtained starting from *both* exponentially distributed individual pair inter-contact times and contact rates.

An interesting physical intuition can be highlighted that justifies the above results. Recall that the inter-contact time aggregate is a mixture of the individual pair inter-contact times. From a physical standpoint, power-law aggregates means that some inter-contact times in the mixture can take extremely large values, possibly diverging. Intuitively, such a behaviour can therefore be generated irrespective of the distribution of individual pair inter-contact times, by including in the mixture individual pairs whose contact rate is extremely small, arbitrarily close to 0. This is exactly the effect of drawing rates from Gamma or exponential distributions, which can admit values of the rates arbitrarily close to 0. The same physical intuition is also confirmed by other results we present in Section 5.2 and 6.3.

The final result we present in this section shows under which conditions aggregate inter-contact times follow an exponential distribution, i.e., $F(x) = e^{-\mu x}$. This is shown in Theorem 3.

Theorem 3: When individual pair inter-contact times are exponentially distributed, aggregate inter-contact times are distributed according to an exponential distribution with rate $\mu > 0$ iff the network is homogeneous, i.e. iff all individual pair inter-contact times are exponentially distributed with rate μ:

$$F(x) = e^{-\mu x} \iff f(\lambda) = \delta(\lambda - \mu) \ , \tag{9}$$

where $\delta(\cdot)$ is the Dirac delta function.

Proof: See Appendix B. □

Theorem 3 shows that, with exponential individual inter-contact times, the only case where the aggregate is also exponential is that of a homogeneous network.

5.2 Asymptotic behaviour of aggregate inter-contact time distributions

In this section we present a further set of results derived when rates are drawn from Pareto distributions. For this set of results we are not able to obtain sufficient *and* necessary conditions for obtaining a given aggregate distribution. However, we are still able to show interesting *sufficient* conditions for obtaining aggregate distributions that asymptotically decay as a power-law with or without exponential cutoff. These results are quite interesting, as several papers in the literature have observed aggregate distributions whose tail decays as a power-law with exponential cutoff. Note that studying the asymptotic behaviour is relevant, as it is the tail of the inter-contact time distributions that determine the convergence properties of naïve forwarding protocols [6].

Firstly, we assume that contact rates are distributed according to a Pareto distribution whose CCDF is $F(\lambda) = \left(\frac{k}{\lambda}\right)^{\gamma}, \lambda > k$, and derive the asymptotic behaviour of $F(x)$ for large x. Note that in this case rates are drawn from a Pareto distribution that does not admit values arbitrarily close to 0. Theorem 4 provides the expression for $F(x)$.

Theorem 4: When individual pair inter-contact times are exponentially distributed and rates are drawn from a Pareto distribution whose CCDF is $F(\lambda) = \left(\frac{k}{\lambda}\right)^{\gamma}, \lambda > k$, the tail of the aggregate inter-contact times decays as a power-law with exponential cutoff, i.e.:

$$F(\lambda) = \left(\frac{k}{\lambda}\right)^{\gamma}, \lambda > k \Rightarrow F(x) \sim \frac{e^{-kx}}{kx} \text{ for large } x \tag{10}$$

Proof: See Appendix B. □

Two interesting insights can be drawn from Theorem 4. First, an aggregate distribution whose tail decays as a power-law with exponential cutoff can emerge also when individual pair inter-contact times are exponential. Again, this challenges common hypotheses used in the literature, that assume individual inter-contact times are power-law with exponential cutoff *because* aggregate inter-contact times are distributed according to this law. Second, this result confirms our intuition about the fact that a key reason for an aggregate distributions with a heavy tail is the existence of individual pairs with contact rates arbitrarily close to 0. In the case considered by Theorem 4 this is not possible, and indeed the tail of the aggregate inter-contact times decays faster than a power-law.

We then study the asymptotic behaviour of the aggregate distribution when contact rates are drawn from a Pareto distribution in the form $F(\lambda) = \left(\frac{k}{k+\lambda}\right)^{\gamma}, \lambda > 0$. The following theorem holds.

Theorem 5: When individual pair inter-contact times are exponentially distributed and rates are drawn from a Pareto distribution whose CCDF is $F(\lambda) = \left(\frac{k}{k+\lambda}\right)^{\gamma}, \lambda > 0$, the tail of the aggregate inter-contact times decays as a power-law with shape equal to 2, i.e.:

$$F(\lambda) = \left(\frac{k}{k+\lambda}\right)^{\gamma}, \lambda > 0 \Rightarrow F(x) \sim \frac{1}{x^2} \textit{ for large } x \quad (11)$$

Proof: See Appendix B. □

Theorem 5 confirms once more that the presence of individual pairs with contact rates arbitrarily close to 0 results in heavy tailed aggregate inter-contact times. Again, it also confirms that the presence of significantly heavy tails (shape equal to 2) in the aggregate inter-contact time distribution is not necessarily an indication that individual pair distributions also present a power-law.

5.3 Validation

In this section we validate the results presented before, by comparing the analytical results with simulations. In our simulation model we consider a network of P=100 pairs. The type of distribution of individual inter-contact times is a parameter of the simulator, set to exponential for the results in this section. Rates are drawn at the beginning of each simulation run according to the specific distribution $f(\lambda)$ we want to test. Each simulation run is built as follows. For each pair we generate at least 100 inter-contact times. Specifically, each simulation run reproduces an observation of the network for a time interval T, defined according to the following algorithm. For each pair, we first generate 100 inter-contact times, and then compute the total observation time after 100 inter-contact times, T_p, as the sum of the pair inter-contact times. T is defined as the maximum of $T_p, p = 1, \ldots, P$. To guarantee that all pairs are observed for the same amount of time, we generate additional inter-contact times for each pair until T_p reaches T. In this way we generate at least $100 * 100$ samples of the aggregate inter-contact time distribution (in practice, we have many more samples in each run). From each run we obtain the percentiles of the aggregate distribution indicated in the following plots. We replicate simulation runs at least 30 times with iid seeds, and finally compute the confidence intervals for the percentiles with 99% confidence level. Although often hardly visible, confidence intervals are shown in the plots for all percentiles.

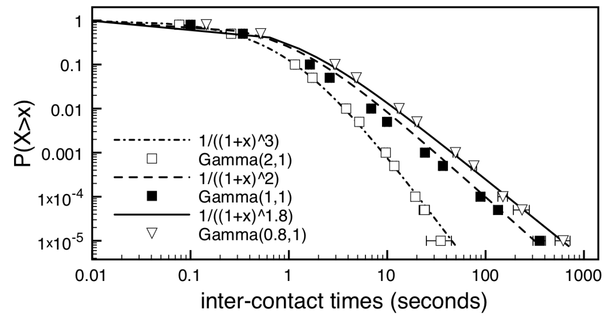

Fig. 2. $F(x)$, contact rates $\Lambda \sim \Gamma(\alpha, b)$

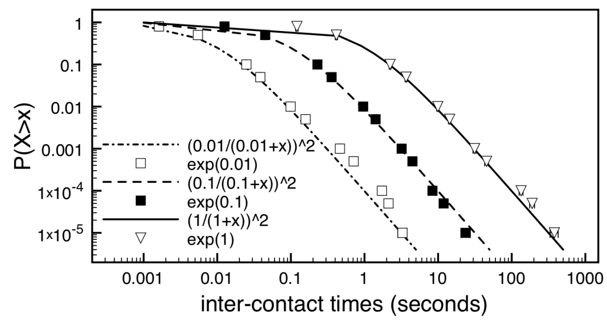

Fig. 3. $F(x)$, contact rates $\Lambda \sim Exp(b)$

Figure 2 shows the aggregate inter-contact times CCDF $F(x)$ when contact rates are drawn from a Gamma distribution with shape equal to 0.8, 1 and 2 (inter-contact times are reported on the x-axis in seconds). According to Theorem 2, this results in aggregate inter-contact times distributed according to a Pareto law with shape $\alpha = 1.8$, 2 and 3, respectively. It is clear that simulation and analytical results are in very good agreement. Figure 3 shows $F(x)$ when the contact rates are exponentially distributed with rate $0.01s^{-1}$, $0.1s^{-1}$ and $1s^{-1}$. Also in this case, according to Corollary 1, the aggregate inter-contact times follow Pareto distributions with shape $\alpha = 2$ and scale 0.01, 0.1 and 1, respectively. Figure 3 shows that also in this case analytical results are very well aligned with simulations.

Finally, Figure 4 and 5 show $F(x)$ when the pairs

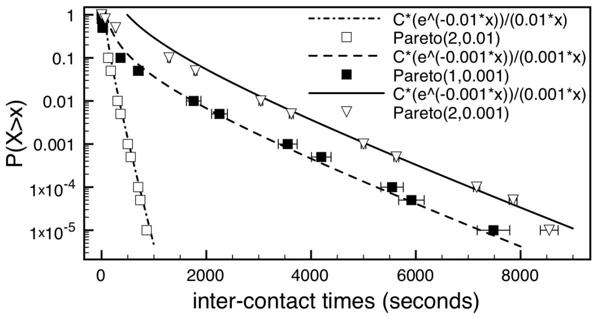

Fig. 4. $F(x)$, contact rates $\Lambda \sim Pareto(\gamma, k)$

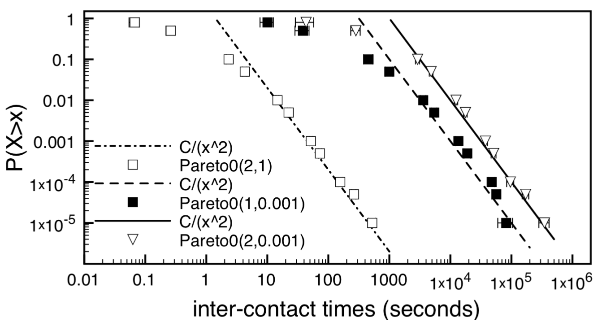

Fig. 5. $F(x)$, contact rates $\Lambda \sim Pareto0(\gamma, k)$

rates are distributed according to a Pareto law $F(\lambda) = \left(\frac{k}{x}\right)^{\gamma}, \lambda > k$ and $F(\lambda) = \left(\frac{k}{k+x}\right)^{\gamma}, \lambda > 0$, respectively (here and in the other Figures in the paper, "C" in the legends of the plots represents a multiplicative constant). From Theorems 4 and 5, the key difference is the fact that in the former case rates cannot be arbitrarily close to 0, while in the latter case they can. The effect on $F(x)$ is to generate a light tail decaying as $\frac{e^{-kx}}{kx}$ in the former case, and a heavy tail decaying as $1/x^2$ in the latter. Recall that in these cases the analysis is not able to capture the complete distribution of $F(x)$, but only its asymptotic behaviour for large x. Figures 4 and 5 confirm that also in this case analytical and simulation results are aligned.

6 NETWORKS WITH POWER-LAW INDIVIDUAL INTER-CONTACT TIMES

In this section we use Theorem 1 to study the dependence between the aggregate and the individual pair inter-contact times when the latter follow different types of power-law distributions. Specifically, in Section 6.1 we analyse the case where individual inter-contacts time follow a Pareto distribution such that inter-contact times arbitrarily close to 0 are possible ("Pareto0" distributions). Section 6.2 presents the case where indivudal inter-contact times follow a Pareto distribution which does not allow values arbitrarily close to 0. In both sections individual inter-contact times present a heavy tail. Finally, in Section 6.3 we consider inter-contact times following a power-law with exponential cutoff distribution, which therefore does not present a heavy tail. Note that, unlike in the exponential case, it is possible to derive closed form analytical results only for the asymtotic behaviour of the aggregate inter-contact time distribution, i.e. for large values of x.

6.1 "Pareto0" individual inter-contact times

We consider the case where the CCDF of individual inter-contact times is in the form

$$F_{\lambda}(x) = \left(\frac{q}{q+x}\right)^{\eta}, \ \eta > 0, \ q > 0 \ x > 0. \qquad (12)$$

To study the CCDF of the resulting aggregate inter-contact time distribution according to Theorem 1, it is necessary to substitute Equation 12 in Equation 5. Remember from Section 3.2 that in our model, for each individual pair p, the contact rate λ_p (i.e., the reciprocal of the average inter-contact time) is sampled from a r.v. with density $f(\lambda)$. For each individual inter-contact time distribution, there is, therefore, a dependence between the parameters $\{q, \eta\}$ and λ_p that must be made explicit before replacing Equation 12 in Equation 5. The only condition that can be imposed is that the average inter-contact time is equal to $1/\lambda_p$, i.e.

$$E[X_p|\lambda_p] = \frac{q}{\eta - 1} \triangleq \frac{1}{\lambda_p} \ , \qquad (13)$$

where the r.v. X_p denotes the inter-contact times of pair p. As we have only one condition to determine two parameters $\{q, \eta\}$, we need to impose one more condition. In the following we consider a natural choice, i.e. we assume that one of the two parameters is fixed, and thus the specific values of the contact rate λ_p impact on the other parameter.

We start by fixing the shape parameter of the Pareto distribution, η. Note that, as the coefficient of variation of a Pareto0 distribution is $\sqrt{\frac{\eta}{\eta-2}}$, fixing the shape of the Pareto0 distributions means fixing the coefficient of variation. We obtain the result in Theorem 6.

Theorem 6: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_{\lambda}(x) = \left(\frac{q}{q+x}\right)^{\eta}$ and the shape parameter η is the same across all pairs, irrespective of the distribution of contact rates, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law with exponent η, i.e. $F(x) \simeq x^{-\eta}$, provided $\eta > 1$ and the following condition holds true:

$$\int_0^{\infty} \lambda f(\lambda) \left(\frac{\eta - 1}{\lambda}\right)^{\eta} d\lambda < \infty \ , \qquad (14)$$

where $f(\lambda)$ is the density of the contact rate distribution.

Proof: See Appendix C. □

The result in Theorem 6 tells that, no matter how contact rates are distributed, provided the integral in Equation 14 converges, when individual inter-contact times follow a Pareto0 distribution with the same shape parameter, also the aggregate distribution presents a heavy tail, with exactly the same exponent. This is clearly a case where the aggregate distribution is representative of the individual pair distributions, at least as far as their behaviour for large x. Thus, using the aggregate distribution to study the convergence properties of forwarding protocols is correct in this case.

We now consider the case where the scale parameter q of the Pareto0 distribution is fixed, and the shape η varies with the contact rate λ. In this case we are not able to obtain general analytical results for any distribution of the contact rates, as in Theorem 6. However, it is still possible to derive analytical results for the specific contact rate distributions that we have considered in the paper, i.e. Gamma, Pareto and Pareto0. Specifically, the following Theorem holds.

Theorem 7: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_\lambda(x) = \left(\frac{q}{q+x}\right)^\eta$ and the scale parameter q is the same across all pairs, if contact rates follow a Gamma, Pareto0, or Pareto distribution, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law. Specifically, the following holds true:

- if contact rates follow a Gamma distribution $\Gamma(\alpha, b)$ then $F(x) \simeq \frac{C}{x(\ln x)^{\alpha+1}}$ holds true for large x, C being a constant greater than 0. Moreover, it can also be shown that $\lim_{x\to\infty} F(x) > \frac{C}{x^{1+\beta}}$, for any $\beta > 0$;
- if contact rates follow a Pareto0 distribution $Pareto0(\gamma, k)$ then $F(x) \simeq \frac{C}{xg(x)}$ holds true for large x, C being a constant greater than 0 and $g(x)$ being a function that, for large x, goes to 0 more slowly than $x^{-\beta}$ for any $\beta > 0$. Therefore, $\lim_{x\to\infty} F(x) > \frac{C}{x^{1+\beta}}$ holds true for any $\beta > 0$;
- if contact rates follow a Pareto distribution $Pareto(\gamma, k)$ then $F(x) \simeq \frac{C}{x^{kq+1}\ln x}$ for large x, C being a constant greater than 0. Therefore, $\lim_{x\to\infty} F_\lambda(x) > \frac{C}{x^{1+kq+\beta}}$ holds true for any $\beta > 0$.

Proof: See Appendix C. □

Theorem 7 shows that, for Gamma, Pareto0 and Pareto contact rates, if the individual inter-contact times follow a Pareto0 distribution also the distribution of the aggregate inter-contact times presents a heavy tail. In particular, for contact rates following a Gamma and Pareto0 distribution, the tail of the aggregate distribution of inter-contact times can be lower bounded by power laws with an exponent arbitrarily close to 1, which is an indication of a particularly heavy tail. Note that in these cases, although the aggregate distribution is power law as the individual pair distributions, no indications can be obtained from the aggregate distribution to assess possible divergence of forwarding protocols, because the aggregate distribution does not provide any information about the shapes of the individual pair distributions.

6.2 "Pareto" individual inter-contact times

In this case the individual inter-contact times follow a Pareto distribution whose CCDF is

$$F_\lambda(x) = \left(\frac{q}{x}\right)^\eta, \ \eta > 0, \ q > 0, \ x > q. \tag{15}$$

We follow the same approach of Section 6.1, by fixing the shape (scale) parameter and letting the scale (shape) parameter vary with the average inter-contact time $1/\lambda$. If we fix the shape parameter η, we obtain a result similar to that of Theorem 6.

Theorem 8: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_\lambda(x) = \left(\frac{q}{x}\right)^\eta$ and the shape parameter η is the same across all pairs, irrespective of the distribution of contact rates, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law with exponent η, i.e. $F(x) \simeq x^{-\eta}$, provided $\eta > 1$ and the following condition holds true:

$$\int_0^\infty \lambda f(\lambda) \left(\frac{\eta-1}{\lambda\eta}\right)^\eta d\lambda < \infty \ , \tag{16}$$

where $f(\lambda)$ is the density of the contact rate distribution.

Proof: See Appendix C. □

Also in this case, the distribution of aggregate inter-contact times presents a heavy tail with the same exponent of the shape of individual pairs, for any contact rate distribution such that the integral in Equation 16 converges. Therefore, the distribution of aggregate inter-contact times is representative of the distributions of individual pairs, as far as convergence of forwarding protocols is concerned. Finally, note that also in this case fixing the shape parameter means assuming that the coefficient of variation of all individual inter-contact times is the same, as it is equal to $\sqrt{\frac{1}{\eta(\eta-2)}}$.

Unlike in the case of Pareto0 individual inter-contact times, when we fix the scale parameter, we are not able to obtain closed form expressions for the distribution of aggregate inter-contact times, even for specific distributions of contact rates, and only numerical solutions can be found.

6.3 "Pareto with cutoff" individual inter-contact times

In this section we consider individual inter-contact times following Pareto distributions with a exponential cutoff in the tail. Denoting with η and q the shape and scale parameters of the Pareto part, and with μ the rate of the exponential part of the distribution, respectively, the CCDF is as follows [18]:

$$F_\lambda(x) = \frac{\Gamma(1-\eta, \mu x)}{\Gamma(1-\eta, \mu q)} \ \eta > 1, \ \mu, q > 0 \ , \tag{17}$$

where $\Gamma(s,x) = \int_x^\infty t^{s-1} e^{-s} dt$ is the upper incomplete Gamma function. Also in this case, to study the properties of the distribution of aggregate inter-contact times (thorough Equation 5), it is first necessary to make explicit the dependence between the parameters of the distributions of individual inter-contact times (η, q, and μ), and the average inter-contact time of the generic pair p, by imposing that $E[X_p|\lambda_p]$ be equal to $1/\lambda_p$. From Equation 17 we obtain

$$E[X_p|\lambda_p] = \frac{1-\eta}{\mu} + \frac{(\mu q)^{1-\eta} e^{-\mu q}}{\mu \Gamma(1-\eta, \mu q)} \triangleq \frac{1}{\lambda_p} \ . \quad (18)$$

In general it is not possible from Equation 18 to find closed forms to make explicit the dependence of η, q and μ on λ_p, as the three parameters of the distribution of individual inter-contact times all appear as parameters of the incomplete Gamma function. However, it is possible to find closed forms for specific cases, where the function $\Gamma(s,x)$ admits exact or approximate closed forms. Recalling that $\eta > 1$ must hold, the only such cases are where the second parameter of Γ either is 0 or tends to ∞. Considering the semantic of the parameters η, q and μ, the only meaningful cases are $\mu \to \infty$ and $q \to 0$. The first case corresponds to a very quick decay of the exponential tail, while the second one corresponds to the possibility of inter-contact times very close to 0. We analyse these two cases separately in the following sections.

6.3.1 Very large rates

When $\mu \to \infty$ the quantity $\Gamma(1-\eta, \mu q)$ can be approximated as $(\mu q)^{-\eta} e^{-\mu q}$ [19]. Therefore, Equation 18 becomes

$$E[X_p|\lambda_p] \simeq \frac{1-\eta}{\mu} + q \simeq q = \frac{1}{\lambda_p} \ . \quad (19)$$

Equation 19 immediately shows the dependence between q and λ_p. In particular, it tells that the case where q is fixed across all pairs is not that interesting, because it corresponds to a homogeneous network where all pairs meet with the same contact rate (equal to $1/q$), and thus the distributions of the aggregate and individual inter-contact times are exactly the same. On the other hand, Equation 19 does not provide any indication on the dependence between η and λ_p. We thus consider again the case where η is fixed across all pairs (as we did in Sections 6.1 and 6.2). Under these conditions, the following theorem holds true.

Theorem 9: When individual pair inter-contact times follow a Pareto distribution with exponential cutoff with shape, scale and rate parameters η, q and μ, if μ is very large and η is the same across all pairs, then the CCDF of the aggregate inter-contact times $F(x)$ decays, for large x, as a Pareto distribution with exponential cutoff with the same shape and rate parameters η and μ, i.e. $F(x) \simeq (\mu x)^{-\eta} e^{-\mu x}$, provided the following condition holds true

$$\int_0^\infty \frac{\lambda f(\lambda)}{\Gamma(1-\eta, \frac{\mu}{\lambda})} d\lambda < \infty \ . \quad (20)$$

Proof: See Appendix C. □

Theorem 9 shows another case where the distribution of aggregate inter-contact times is representative of the distributions of individual pairs, irrespective of the type of network heterogeneity (i.e. of the contact rate distribution). Note that the integral diverges for contact rates following a Gamma or Pareto0 distribution, for any $\mu > 1$, while it admits numerical solutions for Pareto contact rates. This is aligned with the indication we have obtained several times, that contact rate distributions allowing values arbitrarily close to 0 result in power law aggregate inter-contact time distributions. In fact, for Gamma and Pareto0 contact rates, the result in Theorem 9, which predicts a light tail, does not apply.

6.3.2 Very small scales

When $q \to 0$ the quantity $\Gamma(1-\eta, \mu q)$ becomes the constant $\Gamma(1-\eta)$, and thus Equation 18 simplifies as follows:

$$E[X_p|\lambda_p] \simeq \frac{1-\eta}{\mu} \triangleq \frac{1}{\lambda_p} \ . \quad (21)$$

We use again the approach of fixing one among μ or η to study the properties of the resulting distribution of aggregate inter-contact times. In the former case (fixed μ), no closed form expressions have been found, even for the specific distributions of contact rates considered throughout the paper (Gamma, Pareto0 and Pareto). On the other hand, when η is fixed across all pairs, it is possible to find closed form expressions when the contact rates follow a Gamma or a Pareto0 distribution (no closed form expressions have been found in the Pareto case). Specifically, the following theorem holds.

Theorem 10: When individual inter-contact times follow a Pareto distribution with exponential cutoff, whose scale parameter tends to 0 and whose shape parameter is fixed across all pairs, the distribution of aggregate inter-contact times $F(x)$ presents, for large x, a heavy tail, provided $\eta \in (0,1)$ holds true. Specifically:

- if contact rates follow a Gamma distribution $\Gamma(\alpha, b)$ then $\lim_{x\to\infty} F(x) = Cx^{-(\alpha+1)}$, C being a constant greater than 0;
- if contact rates follow a Pareto0 distribution then $\lim_{x\to\infty} F(x) = Cx^{-2}$, C being a constant greater than 0;

Proof: See Appendix C. □

The result in Theorem 10 is particularly interesting as it shows another case where, even though the individual pair inter-contact times do not present a heavy tail, the distribution of aggregate inter-contact times does present a heavy tail. Note that this can be

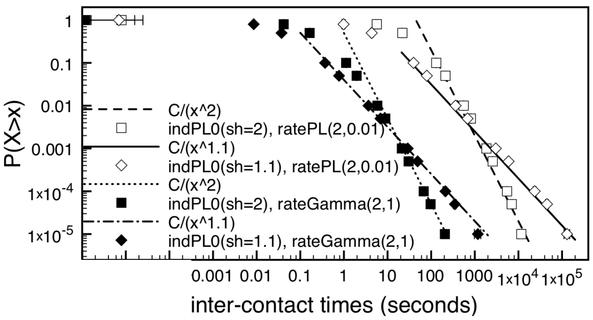

Fig. 6. $F(x)$, individual ICT $X \sim Pareto0$ with fixed shape. Contact rates are $Pareto(2, 0.01)$ or $\Gamma(2, 1)$.

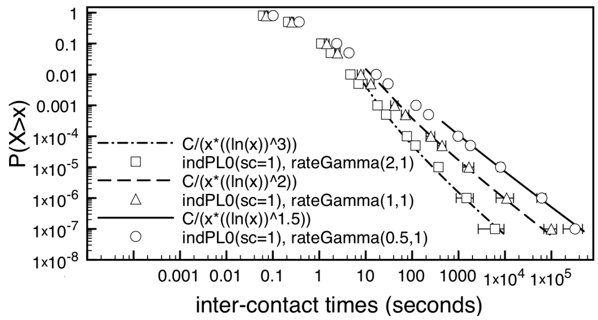

Fig. 8. $F(x)$, individual ICT $X \sim Pareto0$ with fixed scale. Contact rates are Γ.

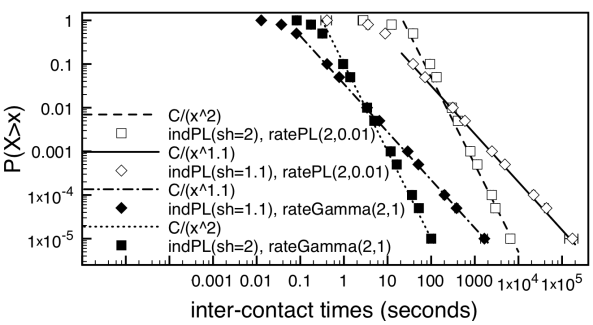

Fig. 7. $F(x)$, individual ICT $X \sim Pareto$ with fixed shape. Contact rates are $Pareto(2, 0.01)$ or $\Gamma(2, 1)$.

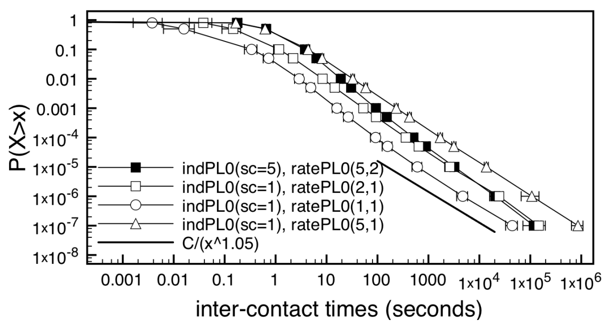

Fig. 9. $F(x)$, individual ICT $X \sim Pareto0$ with fixed scale. Contact rates are $Pareto0$.

proven for contact rates that admits values arbitrarily close to 0, such as rates following a Gamma or a Pareto0 distribution.

6.4 Validation

In this section we compare analytical and simulation results for the Theorems presented in Sections 6.1, 6.2 and 6.3. The simulation model and methodology are the same described in Section 5.3. Firstly, we present results for those cases where the aggregate inter-contact time distribution is representative of the individual pair distributions, i.e. for individual inter-contact times following a Pareto0 or Pareto distribution with fixed shape parameter (Theorems 6 and 8, respectively). Due to practical reasons in obtaining simulation results, we don't present results for the case where individual inter-contact times are power law with exponential cutoff (Theorem 9). Recall that in this case the aggregate distribution decays as $F(x) \simeq (\mu x)^{-\eta} e^{-\mu x}$, and the result holds true for *large* μ. In these conditions, the tail of the aggregate distribution decays so fast that, in simulation, it becomes practically impossible to distinguish between the different percentiles. To obtain significant results from simulation, it is necessary to consider a range for μ where the results of Theorem 9 do not hold anymore.

In the Figures we present hereafter, the legends of the simulation plots have the general form `indXXX(<par>), rateYYY(<par>)`, where `indXXX(<par>)` denotes the distributions of the individual inter-contact times, and `rateYYY(<par>)` the distribution of the contact rates. `XXX(<par>)` and `YYY(<par>)` are replaced in each case by the specific distributions and parameters. `PL`, `PL0`, `PL-CO`, `Gamma` denote Pareto, Pareto0, Pareto with exponential cutoff, and Gamma distributions, respectively. For example, in Figure 6 we plot cases when the individual inter-contact times are Pareto0 with fixed shape equal to 1.1 and 2, while rates are either Pareto (with shape 2 and scale 0.01) or Gamma (with shape 2 and scale 1). This corresponds to the strings `indPL0(sh=1.1)`, `indPL0(sh=2)`, `ratePL(2,0.01)`, `rateGamma(2,1)` which are combined to form the indicated legends. The same convention is also used in the other Figures.

Figures 6 and 7 confirm the results of Theorems 6 and 8. In case of Pareto0 and Pareto individual pair distributions with fixed shape parameter, the aggregate distribution is power law with the same exponent. This holds true for different rate distributions, which do not play any specific role, other than defining a multiplicative constant for $F(x)$.

Figures 8, 9 and 10 confirm the results of Theorem 7, which analyses the case of Pareto0 individual pair distributions with fixed scale. Recall that for rates following a Γ and Pareto distribution, we are able to derive closed form expressions for the tail behaviour, which match very well simulation results. Specifically, note that, for Γ rates (Figure 8), while the shape of the

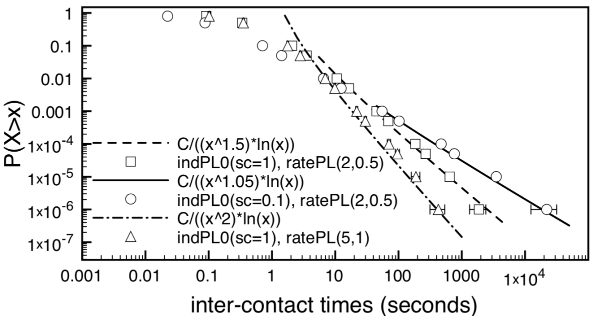

Fig. 10. $F(x)$, individual ICT $X \sim Pareto0$ with fixed scale. Contact rates are $Pareto$.

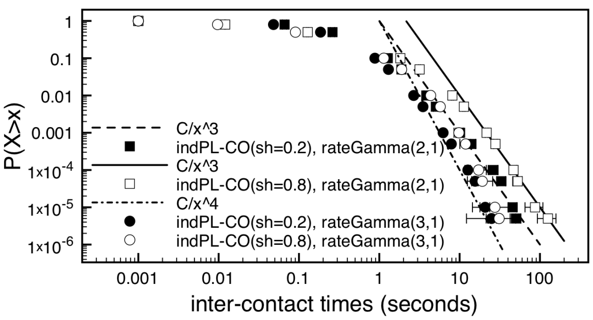

Fig. 11. $F(x)$, individual ICT $X \sim Pareto\text{-}CO$ with small scale and fixed shape. Contact rates are Γ.

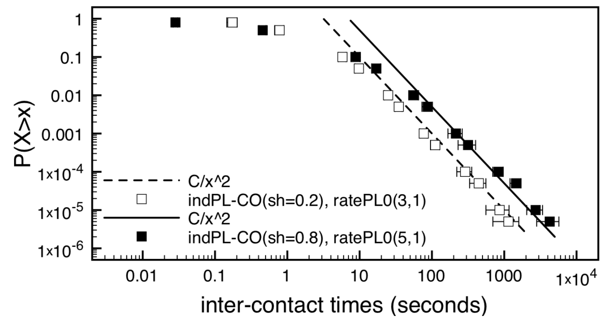

Fig. 12. $F(x)$, individual ICT $X \sim Pareto\text{-}CO$ with small scale and fixed shape. Contact rates are $Pareto0$.

tail initially depends on the parameters of the rate distribution, this dependence disappears for very large x, as $F(x)$ can be lower bounded by $x^{-(1+\beta)}$, $\beta > 0$. The same asymptotic behaviour holds also for Pareto0 rates (Figure 9). In this case, however, we are not able to obtain as precise closed form expressions for the tail behaviour. Recall that $F(x)$ can be approximated, for large x, as $1/(xg(x))$ where $g(x)$ is a function that goes to 0 more slowly than $x^{-\beta}$, $\beta > 0$. As $g(x)$ does not admit closed form expressions, we are able to derive only the asymptotic behaviour, $F(x) \simeq x^{-(1+\beta)}$, which is confirmed by the plots. Note that these figures confirm, once more, the impact of rate distributions admitting values arbitrarily close to 0 on the aggregate distribution. In these cases (Γ and Pareto0, corresponding to Figures 8 and 9) the parameters of the rate distribution and even those of the individual inter-contact time distributions do not matter anymore in the tail behaviour of the aggregate distribution, which approximates a power law with exponent 1. On the other hand, when rates cannot be arbitrarily close to 0 (e.g., Pareto, Figure 10) the tail behaviour of the aggregate distribution does depend on the parameters of the various distributions (in the considered case, on both scale parameters).

Finally, Figures 11 and 12 confirm the results in Theorem 10, which analyses the case of individual inter-contact times following a Pareto distribution with exponential cutoff. In the case where rates follow a Γ distribution (Figure 11), the shape parameter of the Γ distribution determines the shape of the aggregate distribution (recall that $F(x)$ can be approximated as $x^{-(\alpha+1)}$ for large x, α being the shape of the Γ distribution of the rates). On the other hand, when rates follow a Pareto0 distribution (Figure 12), the aggregate decays as a power law with a shape equal to 2 irrespective of the Pareto0 parameters. Both results are clearly confirmed in the figures. Note, in particular, the quite different set of parameters for the parameters of the various distributions shown in Figure 12, and the fact that they do not impact at all on the shape of the aggregate distribution, as predicted by Theorem 10.

7 Discussion and Conclusions

In this paper we have characterised through an analytical model the dependence between the distributions of individual pair inter-contact times and the resulting aggregate distribution in heterogeneous opportunistic networks (i.e., in networks where the contact patterns between pairs are not iid). In our model individual pair distributions are assumed of the same type, but their parameters can vary from pair to pair. We use the contact rates (the reciprocal of the average inter-contact times) to control the parameters of each pair. Specifically, we assume that contact rates are also a random variable, following a given distribution. The value taken by the contact rate r.v. for a given pair determines the values of the parameters of the pair inter-contact time distribution. Therefore, the distribution of contact rates determines, in our model, the heterogeneity of the network.

Understanding the dependence between individual pair and aggregate inter-contact time distributions is an important subject of investigation. Previous foundational results have clearly shown the impact of the distributions of individual inter-contact times on the performance of forwarding protocols. However, the aggregate distribution is a much more convenient figure to describe opportunistic networks with respect to all the distributions of all individual pairs. Therefore, the former has often been used in the literature, assuming that it correctly represents the latter. To the best of our knowledge, our work is the first one that precisely analyses the dependence between the two, and provides clear results on when using the aggregate is possible and when it is not.

In the paper we have presented a general model describing the dependence between the individual pair inter-contact time distributions, the contact rate distribution, and the aggregate inter-contact time distribution. In addition, we have exploited the model to derive analytical results for several relevant cases of heterogeneous networks, featuring inter-contact time distributions found by analysing real traces. We have considered different individual pair distributions (exponential, power law with or without cutoff), and different contact rate distributions (Gamma, exponential, power law with or without cutoff), which result in power law aggregate distributions, with or without exponential cutoff. According to the analysis of real traces available in the literature (e.g., [6], [7], [25], [24]), these are among the most relevant cases to investigate.

This analysis allowed us to derive several interesting insights. Firstly, we have highlighted cases in which using the aggregate distribution instead of the individual pair distributions is correct. In particular, when individual pairs follow a power law distribution and the shape parameter is fixed across all pairs, the aggregate distribution presents again *exactly* the same power law behaviour, irrespective of the distribution of the rates. Also, when individual pair distributions are power law with an exponential cutoff, and the shape is fixed, then the aggregate distribution also behaves as a power law with exponential cutoff with *exactly* the same parameters. Another such case occurs when the individual pair and the aggregate distributions are exponential. However, this case is only possible for homogeneous networks, i.e. when all individual pair distributions (and therefore the aggregate distribution as well) are identical.

In several other cases the aggregate distribution does not correctly represent the individual pair distributions. This occurs consistently over different individual pair and rate distributions. A common trait in these cases is the significant impact of the rate distribution on the aggregate distribution. When the rate distribution allows rates arbitrarily close to 0 (such as for Gamma and Pareto0 distributions), then the aggregate presents a heavy tail, whose shape might even not depend on any parameter of the individual pair and rate distributions. When, on the other hand, the rate distribution does not allow values arbitrarily close to 0 (such as for Pareto distributions), the aggregate typically[4] presents a light tail, usually in the form of a Pareto with exponential cutoff law. This has an intuitive explanation. When rates of some pair can be close to 0, some inter-contact times can be extremely long, and this results in a long tail behaviour.

These findings allow us to look back to several results presented in the literature. We have been able to reproduce aggregate distributions of the same type as those found in most of the real traces (power laws with or without exponential cutoff), starting from a number of different types of individual inter-contact time and rate distributions. Thus, our results call for a cautionary perspective on the typical methodology used so far, which mostly consisted in analysing the aggregate distribution assuming it well represents the distributions of individual pairs. We have shown that, while this is possible in some cases, this is not correct in general, and may lead to completely wrong conclusion, e.g. on the convergence properties of opportunistic forwarding protocols.

On the other hand, our results are very good news for the practical feasibility of opportunistic networks. We have shown that aggregate distributions following a power law with even a very heavy tail can be obtained in a range of diverse heterogeneous networks. In several such cases, the individual inter-contact time distributions are *light* tailed, and can follow, e.g., exponential or power law with cutoff distributions. Interestingly, this is consistent with those few works in the literature that have analysed individual inter-contact time distributions in real traces. Our results are thus able to reconcile apparently contradicting results found in the literature, such as the fact that light tailed individual inter-contact times are compatible with heavy tailed aggregate inter-contact times. Moreover, our results suggest that in several real traces, forwarding protocols might indeed *not* diverge (because the individual inter-contact times are light tailed), and the heavy tail that can be seen in the aggregate distribution is actually a side effect of the network heterogeneity.

APPENDIX A
PROOF OF THEOREM 1

In this Appendix we provide the complete proof of Theorem 1.

Theorem 1: In a network where the contact rates are distributed with density $f(\lambda)$, the CCDF of the aggregate inter-contact times is as follows:

$$\bar{F}(x) = \frac{1}{E[\Lambda]} \int_0^\infty \lambda f(\lambda) \bar{F}_\lambda(x) d\lambda \ .$$

Proof: For the reader convenience, let us recall the expression of the aggregate inter-contact time in Equation 4:

$$\bar{F}(x) = \lim_{T\to\infty} \sum_{p=1}^{P} \frac{n_p(T)}{N(T)} \bar{F}_p(x) = \sum_{p=1}^{P} \frac{\theta_p}{\theta} \bar{F}_p(x) \ .$$

4. The only exception we have found is when individual inter-contact times are Pareto0 with fixed scale. In this case the aggregate is also power law, due to the fact that all individual inter-contact times are power law.

With respect to the case of Lemma 1, the contact rate associated with a given pair is not known in advance, but is a random variable with density $f(\lambda)$. The expression of $F(x)$ can thus be derived conditioning to a specific set of rates $\lambda_1, \ldots, \lambda_P$, and applying the law of total probability. We thus obtain

$$\begin{aligned} F(x) &= \int_{\lambda_1} \cdots \int_{\lambda_P} F(x|\lambda_1,\ldots,\lambda_P) f(\lambda_1,\ldots,\lambda_P) d\lambda_1 \ldots d\lambda_P = \\ &= \int_{\lambda_1} \cdots \int_{\lambda_P} \frac{\sum_p \lambda_p F_p(x)}{\sum_p \lambda_p} f(\lambda_1)\ldots f(\lambda_P) d\lambda_1 \ldots d\lambda_P \ , \end{aligned}$$

where we have exploited the fact that rates of individual pair inter-contact times are assumed to be independent. For a sufficiently large number of pairs (large P), we can apply the law of large numbers, and approximate $\sum_p \lambda_p$ as $E[\Lambda]P$. Swapping the integrals and the summations, we further obtain:

$$\begin{aligned} F(x) &= \frac{1}{E[\Lambda]P} \sum_p \int_{\lambda_1} \cdots \int_{\lambda_P} \lambda_p F_p(x) f(\lambda_1)\ldots f(\lambda_P) d\lambda_1 \ldots d\lambda_P = \\ &= \frac{1}{E[\Lambda]P} \sum_p \int_0^\infty \lambda F_\lambda(x) f(\lambda) d\lambda \\ &= \frac{1}{E[\Lambda]} \int_0^\infty \lambda f(\lambda) F_\lambda(x) d\lambda \ , \end{aligned}$$

where we have exploited the assumption that rates of individual pair inter-contact times are identically distributed. □

Appendix B
Proofs of results in Section 5

In this appendix we provide the detailed proofs of the Theorems presented in Section 5.

Theorem 2: When individual pair inter-contact times are exponentially distributed, aggregate inter-contact times are distributed according to a Pareto law with parameters $\alpha > 1$ and $b > 0$ iff the contact rates follow a Gamma distribution $\Gamma(\alpha-1, b)$, i.e.

$$F(x) = \left(\frac{b}{b+x}\right)^\alpha \iff f(\lambda) = \frac{b^{\alpha-1}}{\Gamma(\alpha-1)} \lambda^{\alpha-2} e^{-b\lambda} \ .$$

Proof: Starting from Equation 6, we note that the following holds true:

$$\int_0^\infty \lambda f(\lambda) e^{-\lambda x} d\lambda = -\frac{\partial}{\partial x} \int_0^\infty f(\lambda) e^{-\lambda x} d\lambda \ .$$

We can thus rewrite Equation 6 as

$$-E[\Lambda] \left(\frac{b}{b+x}\right)^\alpha = \frac{\partial}{\partial x} \int_0^\infty f(\lambda) e^{-\lambda x} d\lambda = \frac{\partial}{\partial x} \mathcal{L}_x(f(\lambda)) \ ,$$

where $\mathcal{L}_x(f(\lambda))$ denotes the Laplace transform of $f(\lambda)$. Integrating over x and computing the inverse Laplace transform, we obtain

$$f(\lambda) = E[\Lambda] \frac{b}{\alpha-1} \frac{b^{\alpha-1}}{\Gamma(\alpha-1)} \lambda^{\alpha-2} e^{-b\lambda} \ .$$

Imposing $\int_0^\infty f(\lambda) d\lambda = 1$ we obtain $E[\Lambda] = \frac{\alpha-1}{b}$, and thus the final expression of $f(\lambda)$, showing that Λ is distributed as $\Gamma(\alpha-1, b)$. Note that the average value of $\Gamma(\alpha-1, b)$ is indeed $\frac{\alpha-1}{b}$ which is consistent with the derivation of $E[\Lambda]$. □

Theorem 3: When individual pair inter-contact times are exponentially distributed, aggregate inter-contact times are distributed according to an exponential distribution with rate $\mu > 0$ iff the network is homogeneous, i.e. iff all individual pair inter-contact times are exponentially distributed with rate μ:

$$F(x) = e^{-\mu x} \iff f(\lambda) = \delta(\lambda - \mu) \ ,$$

where $\delta(\cdot)$ is the Dirac delta function.

Proof: The proof follows the same steps of the proof of Theorem 2. As we want the aggregate distribution to be exponential, we can specialise Equation 5 as follows:

$$\frac{1}{E[\Lambda]} \int_0^\infty \lambda f(\lambda) e^{-\lambda x} d\lambda = e^{-\mu x} \qquad (22)$$

Noting, again, that the following holds true

$$\begin{aligned} \int_0^\infty \lambda f(\lambda) e^{-\lambda x} d\lambda &= -\frac{\partial}{\partial x} \int_0^\infty f(\lambda) e^{-\lambda x} d\lambda \\ &= -\frac{\partial}{\partial x} \mathcal{L}_x(f(\lambda)) \ , \end{aligned}$$

we obtain

$$\mathcal{L}_x(f(\lambda)) = \frac{E[\Lambda]}{\mu} e^{-\mu x} \iff f(\lambda) = \frac{E[\Lambda]}{\mu} \delta(\lambda - \mu)$$

where $\delta(\cdot)$ is the Dirac delta function. Imposing the necessary condition for $f(\lambda)$ being a density function we obtain

$$\begin{aligned} 1 = \int_0^\infty f(\lambda) d\lambda \quad &\iff \quad \frac{\mu}{E[\Lambda]} = \int_0^\infty \delta(\lambda - \mu) d\lambda \\ &\iff \quad \mu = E[\Lambda] \end{aligned}$$

This means that the density function $f(\lambda)$ is the Dirac delta function $\delta(\lambda - \mu)$, which is compatible with the condition $E[\Lambda] = \mu$. Therefore, μ is the only possible contact rate, in order for the aggregate distribution being exponential with rate μ. This means that, starting from exponentially distributed individual inter-contact times, we can obtain an exponential aggregate only if the contacst rates are identical and equal to μ, i.e. only if the network is homogeneous. □

Theorem 4: When individual pair inter-contact times are exponentially distributed and rates are drawn from a Pareto distribution whose CCDF is $F(\lambda) = \left(\frac{k}{\lambda}\right)^\gamma, \lambda > k$, the tail of the aggregate inter-contact times decays as a power-law with exponential cutoff, i.e.:

$$F(\lambda) = \left(\frac{k}{\lambda}\right)^\gamma, \lambda > k \Rightarrow F(x) \sim \frac{e^{-kx}}{kx} \ \textit{for large } x$$

Proof: Substituting the expressions of $f(\lambda)$, $E[\Lambda]$ and $F_\lambda(x)$ in Equation 5 we obtain

$$\begin{aligned} F(x) &= \int_k^\infty \frac{\lambda(\gamma-1)}{\gamma k}\frac{\gamma k^\gamma}{\lambda^{\gamma+1}}e^{-\lambda x}d\lambda \\ &= (\gamma-1)(kx)^{\gamma-1}\Gamma(1-\gamma,kx)\ , \end{aligned}$$

where $\Gamma(s,y)$ is the upper incomplete Gamma function. In the limit $x\to\infty$, $\Gamma(s,y)$ can be approximated as $y^{s-1}e^{-y}$ [19]. Therefore, we obtain

$$F(x) \sim (kx)^{\gamma-1}(kx)^{-\gamma}e^{-kx} = \frac{e^{-kx}}{kx} \quad \text{for large } x\ .$$

□

Theorem 5: When individual pair inter-contact times are exponentially distributed and rates are drawn from a Pareto distribution whose CCDF is $F(\lambda) = \left(\frac{k}{k+\lambda}\right)^\gamma, \lambda > 0$, the tail of the aggregate inter-contact times decays as a power-law with shape equal to 2, i.e.:

$$F(\lambda) = \left(\frac{k}{k+\lambda}\right)^\gamma, \lambda > 0 \Rightarrow F(x) \sim \frac{1}{x^2} \textit{ for large } x$$

Proof: Substituting the expressions of $f(\lambda)$, $E[\Lambda]$ and $F_\lambda(x)$ in Equation 5 we obtain

$$\begin{aligned} F(x) &= \int_0^\infty \frac{\lambda(\gamma-1)}{\gamma k}\frac{\gamma k^\gamma}{(k+\lambda)^{\gamma+1}}e^{-\lambda x}d\lambda \\ &= (\gamma-1)\left[e^{kx}(kx+\gamma)(kx)^{\gamma-1}\Gamma(1-\gamma,kx)-1\right] \end{aligned}$$

With respect to the case of Theorem 4, in this case we have to consider higher terms components in the approximation of $\Gamma(s,y)$. Specifically, we use the following approximation, for $x\to\infty$ [19]:

$$\Gamma(s,y) \simeq y^{s-1}e^{-y}\left[1+\frac{s-1}{y}+\frac{(s-1)(s-2)}{y^2}\right] \quad (23)$$

Substituting Equation 23 in the expression of $F(x)$, after simple algebraic manipulations we obtain

$$F(x) \simeq C\left[\frac{\gamma}{(kx)^2}+\frac{\gamma^2(\gamma+1)}{(kx)^3}\right] \simeq \frac{C}{x^2} \quad \text{for large } x\ .$$

□

APPENDIX C
PROOF OF RESULTS IN SECTION 6

In this appendix we provide the detailed proofs of the Theorems presented in Section 6.

Theorem 6: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_\lambda(x) = \left(\frac{q}{q+x}\right)^\eta$ and the shape parameter η is the same across all pairs, irrespective of the distribution of contact rates, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law with exponent η, i.e. $F(x) \simeq x^{-\eta}$, provided $\eta > 1$ and the following condition holds true:

$$\int_0^\infty \lambda f(\lambda)\left(\frac{\eta-1}{\lambda}\right)^\eta d\lambda < \infty\ ,$$

where $f(\lambda)$ is the density of the contact rate distribution.

Proof: For fixed η, from Equation 13 we obtain the expression of the scale parameter as a function of λ, $q = \frac{\eta-1}{\lambda}$. As q must be greater than 0, this results in the condition $\eta > 1$. We can now use the expression of q to compute a closed form of the CCDF of aggregate inter-contact times, by substituting the expression of $F_\lambda(x)$ in Equation 5. We obtain:

$$F(x) = \frac{1}{E[\Lambda]}\int_0^\infty \lambda f(\lambda)\left(\frac{\frac{\eta-1}{\lambda}}{\frac{\eta-1}{\lambda}+x}\right)^\eta d\lambda\ .$$

For large x this can be approximated as

$$\begin{aligned} F(x) &\simeq \frac{1}{E[\Lambda]}\int_0^\infty \lambda f(\lambda)\left(\frac{\eta-1}{x}\right)^\eta d\lambda \\ &= \frac{1}{E[\Lambda]}x^{-\eta}\int_0^\infty \lambda f(\lambda)\left(\frac{\eta-1}{\lambda}\right)^\eta d\lambda\ . \end{aligned}$$

This concludes the proof. □

Theorem 7: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_\lambda(x) = \left(\frac{q}{q+x}\right)^\eta$ and the scale parameter q is the same across all pairs, if contact rates follow a Gamma, Pareto0, or Pareto distribution, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law. Specifically, the following holds true:

- if contact rates follow a Gamma distribution $\Gamma(\alpha,b)$ then $F(x) \simeq \frac{C}{x(\ln x)^{\alpha+1}}$ holds true for large x, C being a constant greater than 0. Moreover, it can also be shown that $\lim_{x\to\infty} F(x) > \frac{C}{x^{1+\beta}}$, for any $\beta > 0$;
- if contact rates follow a Pareto0 distribution $Pareto0(\gamma,k)$ then $F(x) \simeq \frac{C}{xg(x)}$ holds true for large x, C being a constant greater than 0 and $g(x)$ being a function that, for large x, goes to 0 more slowly than $x^{-\beta}$ for any $\beta > 0$. Therefore, $\lim_{x\to\infty} F(x) > \frac{C}{x^{1+\beta}}$ holds true for any $\beta > 0$;
- if contact rates follow a Pareto distribution $Pareto(\gamma,k)$ then $F(x) \simeq \frac{C}{x^{kq+1}\ln x}$ for large x, C being a constant greater than 0. Therefore, $\lim_{x\to\infty} F_\lambda(x) > \frac{C}{x^{1+kq+\beta}}$ holds true for any $\beta > 0$.

Proof: When q is fixed, from Equation 13 we obtain the expression of the shape parameter η as a function of λ, i.e. $\eta = 1+\lambda q$. Using this expression to substitute $F_\lambda(x)$ in Equation 5 we obtain

$$F(x) = \frac{1}{E[\Lambda]}\frac{q}{q+x}\int_0^\infty \lambda f(\lambda)\left(\frac{q}{q+x}\right)^{\lambda q} d\lambda\ . \quad (24)$$

When contact rates follow a Gamma distribution $\Gamma(\alpha,b)$, by replacing the expression of $f(\lambda)$ in Equation 24, we obtain, for large x:

$$F(x) \simeq \frac{C}{x(\ln x)^{\alpha+1}}\ ,$$

C being a constant greater than 0. For any $\beta > 0$ it is true that $\lim_{x\to\infty}\frac{(\ln x)^{\alpha+1}}{x^\beta} = 0$. This means that for large x we can write $\frac{1}{x^\beta} < \frac{1}{(\ln x)^{\alpha+1}}$. Therefore, for large x, $F(x)$ can be lower bounded as follows:

$$F(x) \simeq \frac{C}{x(\ln x)^{\alpha+1}} > \frac{C}{x^{1+\beta}} \; , \beta > 0.$$

This means that the CCDF aggregate inter-contact times decays, for large x, at least as slow as a power law with exponent $1+\beta$.

When contact rates follow a Pareto0 distribution with shape γ and scale k, $F(x)$ becomes

$$\begin{aligned} F(x) &= C\frac{q}{q+x}\int_0^\infty \frac{\lambda}{(k+\lambda)^{\gamma+1}}\left(\frac{q}{q+x}\right)^{\lambda q} d\lambda \\ &\triangleq C\frac{q}{q+x} g(x) \end{aligned}$$

It can be shown that $\lim_{x\to\infty}\frac{x^{-\beta}}{g(x)} = 0$ for any $\beta > 0$. This means that, for large x, $g(x)$ goes to 0 more slowly than $x^{-\beta}$, or, in other words, we can write $g(x) > x^{-\beta}$. Therefore, $F(x)$ can be lower bounded as follows:

$$F(x) \simeq C\frac{q}{q+x} g(x) > \frac{C}{x^{1+\beta}} \; , \beta > 0.$$

Also in this case, therefore, $F(x)$ decays, for large x, at least as slow as a power law with exponent $1+\beta$.

Finally, when contact rates follow a Pareto distribution with shape γ and scale k, $F(x)$ becomes

$$F(x) = \frac{C}{x}\Gamma(1-\gamma, -kq\ln\frac{q}{q+x})\left(-kq\ln\frac{q}{q+x}\right)^{\gamma-1} \; .$$

In the limit $x \to \infty$, by using the usual approximation for the incomplete Gamma function $\Gamma(s,x) \simeq x^{s-1}e^{-x}$ [19], after simple algebraic manipulations we obtain

$$\lim_{x\to\infty} F(x) = \frac{C}{x^{kq+1}\ln x} \; .$$

Noting again that $\lim_{x\to\infty}\frac{lnx}{x^\beta} = 0$ for any $\beta > 0$ we conclude that $F(x)$ can be lower bounded, for large x, as follows:

$$F(x) \simeq \frac{C}{x^{kq+1}\ln x} > \frac{C}{x^{kq+1+\beta}} \; , \beta > 0.$$

This concludes the proof. □

Theorem 8: When individual pair inter-contact times follow a Pareto distribution whose CCDF is in the form $F_\lambda(x) = \left(\frac{q}{x}\right)^\eta$ and the shape parameter η is the same across all pairs, irrespective of the distribution of contact rates, the tail of the distribution of aggregate inter-contact times decays, for large x, as a power law with exponent η, i.e. $F(x) \simeq x^{-\eta}$, provided $\eta > 1$ and the following condition holds true:

$$\int_0^\infty \lambda f(\lambda)\left(\frac{\eta-1}{\lambda\eta}\right)^\eta d\lambda < \infty \; ,$$

where $f(\lambda)$ is the density of the contact rate distribution.

Proof: By fixing the shape parameter η, and recalling that the average value of the individual inter-contact times is, in this case, $E[X|\lambda] = \frac{q\eta}{\eta-1}$ we obtain the expression of q as a function of λ, $q = \frac{\eta-1}{\lambda\eta}$. As q must be greater than 0, we immediately obtain the condition $\eta > 1$. Using the expression of q to substitute $F_\lambda(x)$ in Equation 5 we obtain

$$F(x) = \frac{1}{E[\Lambda]}x^{-\eta}\int_0^\infty \lambda f(\lambda)\left(\frac{\eta-1}{\lambda\eta}\right)^\eta d\lambda \; .$$

This concludes the proof. □

Theorem 9: When individual pair inter-contact times follow a Pareto distribution with exponential cutoff with shape, scale and rate parameters η, q and μ, if μ is very large and η is the same across all pairs, then the CCDF of the aggregate inter-contact times $F(x)$ decays, for large x, as a Pareto distribution with exponential cutoff with the same shape and rate parameters η and μ, i.e. $F(x) \simeq (\mu x)^{-\eta}e^{-\mu x}$, provided the following condition holds true

$$\int_0^\infty \frac{\lambda f(\lambda)}{\Gamma(1-\eta, \frac{\mu}{\lambda})} d\lambda < \infty \; .$$

Proof: Recalling that in this case q can be approximated as $1/\lambda$, and by applying the usual approximation for the upper incomplete Gamma function $\Gamma(s,x) \simeq x^{s-1}e^{-x}$ for large x, we can approximate $F(x)$ as follows:

$$F(x) \simeq \frac{1}{E[\Lambda]}(\mu x)^{-\eta}e^{-\mu x}\int_0^\infty \frac{\lambda f(\lambda)}{\Gamma(1-\eta, \frac{\mu}{\lambda})} d\lambda \; .$$

This concludes the proof. □

Theorem 10: When individual inter-contact times follow a Pareto distribution with exponential cutoff, whose scale parameter q tends to 0 and whose shape parameter η is fixed across all pairs, the distribution of aggregate inter-contact times $F(x)$ presents, for large x, a heavy tail, provided $\eta \in (0,1)$ holds true. Specifically:

- if contact rates follow a Gamma distribution $\Gamma(\alpha, b)$ then $\lim_{x\to\infty} F(x) = Cx^{-(\alpha+1)}$, C being a constant greater than 0;
- if contact rates follow a Pareto0 distribution then $\lim_{x\to\infty} F(x) = Cx^{-2}$, C being a constant greater than 0;

Proof: Denoting again with μ the rate of the exponential part of the Pareto distribution with exponential cutoff, recalling that q tends to 0 and η is fixed, from the expression of $E[X|\lambda]$ in Equation 21, we obtain the condition $\mu = \lambda(1-\eta)$. Note that, as μ must be greater than 0, this results in the condition $\eta \in (0,1)$. Replacing the resulting expression of $F_\lambda(x)$ in Equation 5 we obtain, for large x,

$$F(x) \simeq \frac{1}{E[\Lambda]}\frac{[(1-\eta)x]^{-\eta}}{\Gamma(1-\eta)}\int_0^\infty \lambda^{1-\eta} f(\lambda)e^{-(1-\eta)\lambda x} d\lambda \; ,$$

where we have used the usual approximation for $\Gamma(s,x)$ for large x.

When contact rates follow a Gamma distribution $\Gamma(\alpha, b)$, by substituting the expression of $f(\lambda)$ we immediately obtain, for large x,

$$F(x) \simeq Cx^{-(\alpha+1)} \ , C > 0.$$

On the other hand, when contact rates follow a Pareto0 distribution with shape γ and scale k we obtain, for large x,

$$F(x) \simeq Cx^{-\eta} \int_0^{\infty} \frac{\lambda^{1-\eta}}{(\lambda+k)^{\gamma+1}} e^{-(1-\eta)} \lambda x d\lambda \ .$$

It can be shown that the following property holds true:

$$\lim_{x\to\infty} \frac{F(x)}{x^{-2}} = \frac{C\Gamma(2-\gamma)(1-\eta)^{2-\eta}}{k^{\gamma+1}}$$

This means that $F(x)$ decays, for large x as a power law with exponent equal to 2. This concludes the proof. □

References

[1] L. Pelusi, A. Passarella, and M. Conti, "Opportunistic networking: data forwarding in disconnected mobile ad hoc networks," *IEEE Comm. Mag.*, vol. 44, no. 11, pp. 134 –141, 2006.

[2] V. Borrel, M. H. Ammar, and E. W. Zegura, "Understanding the wireless and mobile network space: a routing-centered classification," in *Proceedings of the second ACM workshop on Challenged networks*, ser. CHANTS '07. New York, NY, USA: ACM, 2007, pp. 11–18. [Online]. Available: http://doi.acm.org/10.1145/1287791.1287795

[3] Y. Chen, V. Borrel, M. Ammar, and E. Zegura, "A framework for characterizing the wireless and mobile network continuum," *SIGCOMM Comput. Commun. Rev.*, vol. 41, pp. 5–13. [Online]. Available: http://doi.acm.org/10.1145/1925861.1925863

[4] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, and H. Weiss, "Delay-tolerant networking: an approach to interplanetary internet," *Communications Magazine, IEEE*, vol. 41, no. 6, pp. 128 – 136, june 2003.

[5] S. Jain, K. Fall, and R. Patra, "Routing in a delay tolerant network," in *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, ser. SIGCOMM '04. New York, NY, USA: ACM, 2004, pp. 145–158. [Online]. Available: http://doi.acm.org/10.1145/1015467.1015484

[6] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott, "Impact of human mobility on opportunistic forwarding algorithms," *IEEE Trans. Mob. Comput.*, vol. 6, no. 6, pp. 606–620, 2007.

[7] T. Karagiannis, J.-Y. L. Boudec, and M. Vojnovic, "Power law and exponential decay of inter contact times between mobile devices," in *MOBICOM*, 2007, pp. 183–194.

[8] ——, "Power law and exponential decay of intercontact times between mobile devices," *IEEE Trans. Mob. Comput.*, vol. 9, no. 10, pp. 1377–1390, 2010.

[9] Z. J. Haas and T. Small, "A new networking model for biological applications of ad hoc sensor networks," *IEEE/ACM Trans. Netw.*, vol. 14, pp. 27–40, February 2006. [Online]. Available: http://dx.doi.org/10.1109/TNET.2005.863461

[10] M. Grossglauser and D. N. C. Tse, "Mobility increases the capacity of ad hoc wireless networks," *IEEE/ACM Trans. Netw.*, vol. 10, pp. 477–486, August 2002. [Online]. Available: http://dx.doi.org/10.1109/TNET.2002.801403

[11] T. Spyropoulos, K. Psounis, and C. S. Raghavendra, "Efficient routing in intermittently connected mobile networks: the multiple-copy case," *IEEE/ACM Trans. Netw.*, vol. 16, pp. 77–90, February 2008. [Online]. Available: http://dx.doi.org/10.1109/TNET.2007.897964

[12] ——, "Efficient routing in intermittently connected mobile networks: the single-copy case," *IEEE/ACM Trans. Netw.*, vol. 16, pp. 63–76, February 2008. [Online]. Available: http://dx.doi.org/10.1109/TNET.2007.897962

[13] A. Vahdat and D. Becker, "Epidemic routing for partially connected ad hoc networks," Duke University, Tech. Rep., 2000.

[14] M. McNett and G. M. Voelker, "Access and mobility of wireless pda users," *ACM MC2R*, vol. 9, 2005.

[15] T. Henderson, D. Kotz, and I. Abyzov, "The changing usage of a mature campus-wide wireless network," in *ACM MOBICOM*, 2004.

[16] J. Su, A. Chin, A. Popivanova, A. Goel, and E. de Lara, "User mobility for opportunistic ad-hoc networking," in *IEEE MSCA*, 2004.

[17] N. Eagle and A. (Sandy) Pentland, "Reality mining: sensing complex social systems," *Personal Ubiquitous Comput.*, vol. 10, pp. 255–268, March 2006. [Online]. Available: http://dx.doi.org/10.1007/s00779-005-0046-3

[18] A. Clauset, C. R. Shalizi, and M. E. J. Newman, "Power-law distributions in empirical data," *SIAM Reviews*, Jun. 2007. [Online]. Available: http://arxiv.org/abs/0706.1062

[19] M. Abramowitz and I. A. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ser. NBS Applied Mathematics Series 55. National Bureau of Standards, 1964.

[20] V. Borrel, F. Legendre, M. D. de Amorim, and S. Fdida, "Simps: using sociology for personal mobility," *IEEE/ACM Trans. Netw.*, vol. 17, no. 3, pp. 831–842, 2009.

[21] H. Cai and D. Y. Eun, "Crossing over the bounded domain: From exponential to power-law intermeeting time in mobile ad hoc networks," *Networking, IEEE/ACM Transactions on*, vol. 17, no. 5, pp. 1578 –1591, oct. 2009.

[22] ——, "Toward stochastic anatomy of inter-meeting time distribution under general mobility models," in *ACM MobiHoc*, 2008, pp. 273–282.

[23] C. Boldrini and A. Passarella, "HCMM: Modelling spatial and temporal properties of human mobility driven by users' social relationships," *Computer Communications*, vol. 33, no. 9, pp. 1056–1074, 2010.

[24] V. Conan, J. Leguay, and T. Friedman, "Characterizing pairwise inter-contact patterns in delay tolerant networks," in *Autonomics*, 2007.

[25] W. Gao, Q. Li, B. Zhao, and G. Cao, "Multicasting in delay tolerant networks: a social network perspective," in *ACM MobiHoc*, 2009.

[26] M. Conti, S. Giordano, M. May, and A. Passarella, "From opportunistic networks to opportunistic computing," *IEEE Communications Magazine*, vol. 48, no. 9, pp. 126–139, Sep. 2010. [Online]. Available: http://dx.doi.org/10.1109/MCOM.2010.5560597

[27] A. Passarella and M. Conti, "Characterising aggregate inter-contact times in heterogeneous opportunistic networks," in *IFIP Networking*, 2011.