Consiglio Nazionale delle Ricerche

Performance modelling of opportunistic forwarding under heterogenous mobility

C. Boldrini, M. Conti, A. Passarella

IIT TR-12/2013

Technical report

Luglio 2013

Istituto di Informatica e Telematica

Performance modelling of opportunistic forwarding under heterogenous mobility

Chiara Boldrini[a,*], Marco Conti[a], Andrea Passarella[a]

[a] *Institute for Informatics and Telematics, National Research Council, Via G.Moruzzi 1, 56124 Pisa, Italy*

Abstract

The Delay Tolerant Networking paradigm aims to enable communications in disconnected environments where traditional protocols would fail. Opportunistic networks are delay tolerant networks whose nodes are typically the users' personal mobile devices. Communications in an opportunistic network rely on the mobility of users: each message is forwarded from node to node, according to a hop-by-hop decision process that selects the node that is better suited for bringing the message closer to its destination. Despite the variety of forwarding protocols that have been proposed in the recent years, there is no reference framework for the performance modelling of opportunistic forwarding. In this paper we start to fill this gap by proposing an analytical model for the first two moments of the delay and the number of hops experienced by messages when delivered in an opportunistic fashion. This model seamlessly integrates both social-aware and social-oblivious single-copy forwarding protocols, as well as different hypotheses for user contact dynamics. More specifically, the model can be solved exactly in the case of exponential and Pareto inter-meeting times, two popular cases emerged from the literature on human mobility analysis. In order to exemplify how the proposed framework can be used, we discuss its application to two case studies with different mobility settings. Finally, we discuss how the framework can be also solved exactly when inter-meeting times follow a hyper-exponential distribution. This case is particularly relevant as hyper-exponential distributions are

*Corresponding author. Phone + 39 050 3153504 Fax + 39 050 3152593

Email addresses: `chiara.boldrini@iit.cnr.it` (Chiara Boldrini), `marco.conti@iit.cnr.it` (Marco Conti), `andrea.passarella@iit.cnr.it` (Andrea Passarella)

Preprint submitted to Computer Communications *March 15, 2013*

able to approximate the large class of high-variance distributions (distributions with coefficient of variation greater than one), which are those more challenging, e.g., from the delay standpoint.

---

1. Introduction

With the advent of powerful and lightweight mobile devices, such as smartphones and tablets, the ubiquitous networking vision is quickly becoming a reality. A further step in the direction of communicating anytime anywhere is represented by the Delay Tolerant Networking paradigm, which enables communications also in disconnected environments. In such conditions, the main requirement of protocols for legacy Mobile Ad Hoc NETworks (MANET), i.e., the presence of an end-to-end path connecting the source and the destination of a message, can hardly be satisfied. Typical delay tolerant networks are, e.g., networks made up of subnetworks connected only by satellite links [1], or networks whose nodes are people moving around with their hand-held devices [2]. The latter case is the scenario considered in this paper. In the literature, such networks have been named Pocket Switched Networks (PSN [3]) or simply *opportunistic networks*, because they opportunistically exploit contacts between users.

In opportunistic networks, messages are dynamically handed over from node to node upon contact, according to the store-carry-and-forward paradigm. Nodes carry messages with them while they move across the network and with their movements they create transmission opportunities that enable communications. Thus, in opportunistic networks the delay accumulated by the messages along the forwarding path critically depends on the way users move. The simplest exploitation of contact opportunities in order to forward messages is represented by Epidemic forwarding [4], which generates and hands over a new copy of the message for each new encounter. The rationale behind this approach is to leverage as many routes to the destination as possible. Unfortunately, this greedy approach suffers from severe resource consumption and tends to overload the network [5]. Smarter strategies as to who to forward and how many copies should be generated have been devised since then. According to the type of information used when making forwarding decisions, these strategies can be classified as partially social-aware [6][7] and fully social-aware [8][9][10]. They leverage information about the users, their contact dynamics, the environment they operate in, the social relationships they share, in order to select one (or a bunch of) *best* next hop. Depending on the number of copies generated for the same message, forwarding protocols can also be classified into single-copy or multi-copy schemes. In the first case, at any time, in the network there is just one copy of the message to be delivered, while in the second case more copies are generated, hoping that at least one of them will eventually reach the destination. While multi-copy strategies have been shown to improve the reliability of delivery, they are typically resource consuming.

Despite the variety of practical forwarding solutions based on different heuristics (such as encounter frequency and sociality metrics) no general framework has been introduced so far for the analysis of opportunistic forwarding protocols in a structured way. Some models exist in the literature (e.g., [11], [12], [7], [13], [14]), but they are specific to the protocols being studied and can hardly be re-used when the protocols are changed. The situation is even worse for social-aware schemes, which, despite their popularity, are typically difficult to model analytically. Moreover, the absence of a general consensus on some fundamental properties of user movement patterns (e.g., the distribution of the inter-meeting times) makes it even more complex to found a model on a solid basis. In fact, the performance of message forwarding closely depends on the users' contact dynamics [15]. From the analysis of real movement traces many hypotheses (e.g., [15], [16], [17], [18], [19], [20], [21]) have been made as to which distribution better describes significant quantities such as the time between consecutive contacts, or the duration of a contact, but without ultimate consensus.

The contribution of this paper is twofold. First, a general framework for the analysis of single-copy forwarding schemes is introduced. This model, based on Markov chains, allows us to compute significant quantities, such as the first and second moments of the number of hops and delay, which characterize the forwarding performance. These moments can then be used to approximate, e.g., the full distribution of the delay and number of hops.Clearly, the full distribution, e.g., of the delay is more informative than just its expectation, as it allows us to analyze , for example, the dependency of the delay on the TTL. This general framework also takes into account social-awareness, which can be incorporated seamlessly into the model. In addition, our framework is independent of specific mobility assumptions, thus it would remain usable even if new insights on the way users move were provided.

The second contribution is the instantiation of the framework in three different mobility scenarios. More specifically, we solve the framework *exactly* in the case of exponential and power law inter-meeting times, which are popular assumptions for inter-meeting times emerged in the literature [18] [15] [22]. In addition, we also provide a complete solution to the framework in the case of hyper-exponentially distributed inter-meeting times. The latter result is particularly significant, since the hyper-exponential distribution can approximate the behavior of a large class of distributions, those having a coefficient of variation greater than 1. The coefficient of variation [23] is defined as the ratio between the standard deviation and the mean, and measures the dispersion of a probability distribution. The higher the coefficient of variation, the more distant a sampled value can be from the mean. High-variance distributions are extremely important in opportunistic networks for two reasons. First, they have often emerged as a plausible hypothesis for inter-meeting times (apart from the power law hypothesis, recently the LogNormal one has also gained a lot of popularity [24]). Second, high-variance distributions can drastically affect the delay experienced by messages, causing the expectation of the delay to diverge in extreme cases [25] [15].

The characteristics of single-copy schemes have been analytically studied in the literature for what concerns social-oblivious strategies [7] [15], but, to the best of our knowledge, the one proposed in this paper is the first general framework that takes into account the social-awareness of the forwarding process. Moreover, results obtained for single-copy schemes are important to multi-copy schemes as well. Consider for example multi-copy schemes in which replication can occur only at the source node. Each copy travels along a path independently of the others. While the delivery from the source node to the first relays is significantly different from a single-copy delivery due to the multi-copy generation, from the first relay to the destination the delay can be approximated using single-copy results. The extension of the framework to the multi-copy case is currently under study.

The paper is structured as follows. In Section 2 we review the state of the art on forwarding protocols and performance modelling for opportunistic networks. In Section 3 we describe the scenario we consider and the assumptions we make, based on which, in Section 4, we define our general modelling framework. After defining in Section 5 our reference forwarding schemes, in Section 6 the general framework is specialized under the assumptions of exponential and power law inter-meeting times. In order to exemplify how the proposed model can be used, we discuss its application to two case studies with different mobility settings in Section 7. Finally, in Section 8 we solve the model exactly in the case of hyper-exponential inter-meeting times, and we discuss how this case can be used to solve the model approximately in the general case of high-variance inter-meeting times.

2. Related Work

2.1. Opportunistic forwarding

According to the type of information that they exploit when making forwarding decisions, forwarding protocols can be classified into social-oblivious, partially social-aware and fully social-aware protocols [8]. In the following we overview some of the most significant protocols for each of these categories. For a more detailed survey, we refer the reader to Al Hanbali et al. [26]. Social-oblivious protocols do not use at all information on the way nodes meet or relate with each other. This is the case of the Epidemic protocol [4], whose strategy is to generate and hand over a new copy of the message to each node encountered, and of the Direct Transmission protocol [27], in which messages can only be delivered to the destination when encountered directly. Their performance is typically poor because either they consume a lot of resources and overload the network (Epidemic [8]) or they are not able to find a path to the destination even when many are available (as shown in Section 7, the Direct Transmission strategy suffers from this problem). For this reason, they are typically used as a baseline for performance evaluation only. More specifically, Epidemic routing provides the minimum possible delay in ideal settings with infinite resources, while Direct Transmission minimizes the number of hops travelled by messages. In order to mitigate the side effects of Epidemic-style forwarding schemes in resource constrained environments, controlled flooding solutions have been proposed. The Spray&Wait protocol [5] (where only L relays are used) and gossiping [28] (where messages are forwarded with probability p upon encounter) are examples of limited flooding, and still can be classified as social-oblivious protocols. Another popular social-oblivious forwarding protocol is the Two Hop scheme [27], in which a message is forwarded by the source node to the first node encountered, which is then allowed only to pass the message directly to the destination. The Two Hop strategy has been shown to guarantee the maximum throughput capacity in a homogeneous network [27]. Despite their appealing simplicity, these social-oblivious protocols just make a random guess on which path towards the destination the message should follow, and thus they are typically very far from being optimal in networks where the presence of humans, with their highly predictable movements, would provide the basis for more accurate forwarding decisions.

Partially social-aware protocols leverage network-level information such as time since the last encounter (FRESH [29], Spray&Focus [5]), frequency of encounters (PROPHET [6]), and total number of encounters [30]. This information is used to predict future meetings between pairs of nodes and thus to select relays that can guarantee a quick delivery according to the heuristic in use. Partially social-aware protocols, however, do not allow for the intentional exploitation of the intrinsic social component in user mobility but only rely on very simple metrics as the ones mentioned above.

Fully social-aware protocols explicitly exploit the social structure of the network of users in order to make forwarding decisions. This is because social-awareness enables the prediction of user encounters [31], which constitute forwarding opportunities. One approach is based on the exploitation of the roles of the nodes in the social graph associated with the network of users. The main idea is that nodes that are more *central* in the social graph are likely to be better forwarders than the other nodes. BUBBLE [9], SimBet [10], and PeopleRank [32] rely on this approach. On the other hand, social context-aware protocols keep track of a variety of information on the environment – *context* – the users live in (e.g., the people they meet, the friends they have, the places they visit) and use this information to quantify the ability of nodes to deliver messages. HiBOp [8] and SocialCast [33] belong to this group.

2.2. Performance modelling

Performance modelling of opportunistic forwarding algorithms has been the subject of several papers. Zhang et al. [11], Haas and Small [28], and Groenevelt et al. [34] focus on the modelling of Epidemic-style routing, either by means of Markov chains or fluid (Ordinary Differential Equations) models. A class of two-hop forwarding schemes is studied by Al Hanbali et al. [12] [35], again relying on Markov chain theory. A variety of single-copy forwarding schemes have been analysed by Spyropoulos et al. [7] by means of random walks on a graph. Their approach shares many similarities with this paper but, analogously to the contributions cited above, it relies on the exponential assumption for node inter-meeting times and assumes a homogeneous network, i.e., all node pairs being i.i.d. from the contact process standpoint. In this paper, instead, we relax these assumptions and we consider both heterogeneous mobility and various distributions for the inter-meeting times. As a matter of fact, homogeneous contact dynamics have been shown to be unrealistic [17]: some users may cluster and move together, others may never get in touch with each other. For this reason, models taking into account node diversity are needed.

To the best of our knowledge, heterogeneous contact patterns have been considered by Spyropoulos et al. [13], Lee and Eun [14], Picu et al. [36] [37], and Ip et al. [38]. The latter, however, only considers two classes of nodes from the mobility standpoint, and focuses only on Epidemic dissemination. Spyropoulos et al. [13] propose a more complete analysis, including multiple classes and a variety of forwarding protocols. However, they still rely on the exponential assumption for inter-meeting times. Lee and Eun [14] study the performance of a class of two-hop forwarding policies under heterogeneous contact dynamics, but the distribution of the inter-meeting times is considered exponential. As this work, [36] [37] aim to model the performance of DTN data delivery strategies in realistic settings. The main difference with this work lies again in the fact that the model proposed in [36] [37] has the exponential assumption for inter-meeting times as its key approximation. Differently, here we explicitly model those cases in which inter-meeting times are different from exponential, providing also a general strategy for the larger class of high-variance distributions.

There are not many contributions that tackle the modelling of opportunistic forwarding relaxing the exponential assumption for inter-meeting times. The only existing works that consider different distributions are those by Chaintreau et al. [15] and Lee and Eun [39]. The latter is focused on capacity scaling issues, which are not studied in this paper. The contribution of Chaintreau et al. [15] is foundational in the field of opportunistic networking. but it is mainly focused on the derivation of conditions on the power law exponent α of inter-meeting times under which forwarding protocols can provide finite expected delay.

This paper extends [40], where a simplified version of the framework discussed here was presented. More specifically, [40] only focused on exponential inter-meeting times, and solved the framework (only in terms of the expectation of the delay and the number of hops) in this case. Here we extend the modeling to the power law and hyper-exponential cases, with the latter providing a general way for approximating the behavior of forwarding protocols when inter-meeting times belong to the large class of high-variance distributions. The work presented here is also complementary to our work in [41], where we have studied the behavior of forwarding schemes under estimation errors on the forwarding parameters. Such estimation errors makes infeasible the kind of exact analysis that we perform in this paper, thus different, approximated, techniques were used in [41].

3. Network Model

We first introduce the network model and the notation (Table 1) that we use throughout the paper.

Our model considers a network with N mobile nodes. For the sake of simplicity, we hereafter assume that messages can be exchanged only at the beginning of a contact between a pair of nodes and that the transmission of the relayed messages can be always completed within the duration of a contact. In addition, we assume that each message is a bundle [1], an atomic unit that cannot be fragmented. We also assume infinite buffer space on nodes. Given that we are considering single-copy schemes, buffer size is not expected to be critical, at least from low to medium network load. All the above assumptions allow us to isolate, and thus focus on, the effects of node mobility from other effects, and are common assumptions in the literature on opportunistic networks modelling (they are used in most of the literature reviewed in Section 2).

Given that messages are handed over from node to node before reaching their destination, the way nodes move heavily affects the delay experienced by messages. As we assume that the transmission of a message can always be completed during a pair-wise contact, the actual duration of the contact is not critical. Thus, the main role in the experienced delay is played by inter-meeting times, which are defined as follows.

Definition 1 (Inter-meeting Time)**.** *The inter-meeting time M_{ij} between node i and node j is defined as the time between two consecutive meetings between the same pair of nodes. If t_f is the time at which a contact between node i and node j has* just *finished, the inter-meeting time M_{ij} is given by:*

$$M_{ij} = \min_{t > t_f} \{ t - t_f : ||X_i(t) - X_j(t)|| < r \} \tag{1}$$

where $X_i(t)$ and $X_j(t)$ denote the position of i and j at time t, and r is the transmission range[1]*.*

[1]Without loss of generality, here we assume a deterministic unit disk graph model for radio propagation. In other words, nodes can communicate only if their current distance is smaller than the transmission range. This is a common assumption in the literature on opportunistic networks. The proposed framework still applies for every other model of radio propagation.

N	number of nodes in the network
f_X, F_X	density and complementary cumulative distribution function (CCDF) of random variable X
M_{ij}	inter-meeting time for the i, j node pair
R_{ij}	residual inter-meeting time for the i, j node pair
μ_{ij}	contact rate for the i, j node pair
$\hat{\mu}_{ij}$	contact rate for the i, j node pair resulting from an online estimation process, e.g., by means of pair-wise exchange of history of encounters
$f^{\varphi}_{i,d}$	fitness of node i as a relay to destination d under forwarding policy φ
p_{ij}	transition probabilities of the forwarding Markov process
$p_{ij}^{forw(\varphi)}$	probability that node i hands over the message to node j upon encounter when forwarding policy φ is in use
T_i	time before node i hands over the message to any other node or, equivalently, time before the forwarding Markov process exits from state i
T_{ij}	time before node i hands over the message to node j or, equivalently, time before the forwarding Markov process goes from state i to state j
D_i	delay of a message generated by node i and addressed to node d
H_i	number of hops travelled by a message generated by node i and addressed to node d
$\mathcal{R}_i$	set comprising all nodes that are potential relays from node i, i.e., $p_{ij}^{forw(\varphi)} > 0$

Table 1: Notation

In the following we denote as μ_{ij} the rate of inter-meeting times of the process of encounters between two nodes i and j. We also assume that the network is stationary, thus inter-meeting rates do not vary with time (i.e., $\mu_{ij}(t) = \mu_{ij}$). By definition, $\mu_{ij} = \frac{1}{E[M_{ij}]}$, where $E[M_{ij}]$ denotes the expectation of the inter-meeting time M_{ij} between node i and node j. As we assume that inter-meeting times between every specific node pair i, j are independent and identically distributed, the meeting process between node i and node j can be modelled as a renewal process [42].

The message generation and process and the mobility process are independent. We also assume that nodes do not keep track of the time since the last encounter with any other node. This means that when a node generates a new message (or it receives a new message to relay), the time since the last encounter with any other node is unknown. For this reason, in our analysis we will often use the concept of residual inter-meeting time.

Definition 2 (Residual Inter-meeting Time)**.** *Assuming that node i and node j are not in contact at time t_r, the residual inter-meeting time $R_{ij}(t)$ between them is given by the time interval between t_r and the first time node i and node j come into each other's range again, i.e.:*

$$R_{ij} = \min_{t>t_r}\{t - t_r : ||X_i(t) - X_j(t)|| < r\}, \quad (2)$$

where $X_i(t)$ and $X_j(t)$ denote the position of i and j at time t, and r is the transmission range.

There has been an intense debate in the research community about the probability distribution that better describes the inter-meeting times between users. Chaintreau et al. [15] found that inter-meeting times could be described by a power law distribution (hypothesis later validated theoretically by [22]). After analysing both the same traces and an additional one, Karagiannis et al. [16] suggested that a power law distribution with a final exponential cut-off could better match the actual shape of the inter-meeting times. According to Gao et al. [18], the same traces support instead the hypothesis of exponentially distributed inter-meeting times. Along these contributions, also other hypotheses have been studied (such as LogNormal [17] and Double Pareto LogNormal [43]). This brief overview suggests the need for a more careful and deeper statistical analysis of contact traces, which is clearly out of the scope of this work. In the following, we restrict our analysis to exponential, power law, and hyper-exponential inter-meeting times,

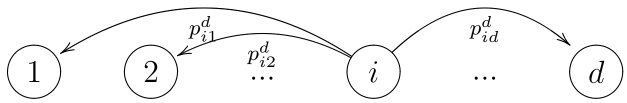

Figure 1: Fragment of the Embedded Markov Chain (valid for all $i \neq d$)

for which we are able to obtain closed form solutions for the parameters of the framework. However, please note that in the other cases, e.g., power law inter-meeting times with exponential cut-off, the model can still be solved numerically.

4. A General Framework for Modelling the Forwarding Process

As discussed above, there is no final agreement on the probability distribution that better describes the inter-meeting process between pairs of mobile users. For this reason, we choose to make our analytical framework as general as possible. Due to its flexibility, we use a semi-Markov process with N states to model the opportunistic forwarding process. A semi-Markov process is one that changes state in accordance with a Markov chain (called *embedded* or *jump* chain) but where transitions between states can take a random amount of time with an arbitrary distribution [42]. As such, it is fully described by the transition matrix associated with its embedded chain and by $T_i, \forall i = 0, \cdots, N$, where T_i denotes the distribution of the time that the semi-Markov process spends in state i before making a transition.

We express our semi-Markov process associated with the single-copy message forwarding process in terms of the embedded Markov chain in Figure 1. Assuming that node i is currently holding a message whose destination[2] is d, the probability p_{ij}^d that node i will delegate the forwarding of the message to another node j is a function of both the likelihood of meeting node j and the probability that node i will hand over the message to node j according to the forwarding policy in use. The transition matrix $\mathbf{T}$ associated with the process of forwarding a message from a source node i to the destination node

[2]The chain is different for different destinations, because the convenient relays are generally not the same. However, for the sake of readability, in the following we drop superscript d

d is given below, where, as an example, $d = N$.

$$\mathbf{T} = \begin{pmatrix} 0 & p_{12} & \dots & p_{1,N-1} & p_{1,N} \\ p_{21} & 0 & \dots & p_{2,N-1} & p_{2,N} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}$$

The state associated with the destination node d is absorbing, because in state d the forwarding process is completed. Please note, however, that there is no guarantee that such absorbing state is eventually reached, due to the potential presence of other closed classes in the forwarding Markov chain.

Once the forwarding Markov process is completely defined in terms of transition probabilities and exit times, we can exploit well known algorithms for Markov chain transient analysis in order to compute significant properties of the forwarding process. More specifically, the delay D_i from node i to a tagged destination d can be written as follows:

$$D_i = \begin{cases} T_{id} & \text{with probability } p_{id} \\ T_{ij} + D_j & \text{with probability } p_{ij}, \forall j \neq d, \end{cases} \tag{3}$$

where T_{ij} denotes the time before node i hands over the message to node j conditioned on the fact that j is the first encountered suitable next hop for node i (corresponding to the time before the chain moves from state i to state j), and p_{ij} is the probability that node j is actually the first encountered suitable next hop for node i. In addition, we have that $D_d = 0$ (the delay from the destination to itself is null). Similarly, the hop count H_i from node i to a tagged destination d can be written as follows:

$$H_i = \begin{cases} 1 & \text{with probability } p_{id} \\ 1 + H_j & \text{with probability } p_{ij}, \forall j \neq d, \end{cases} \tag{4}$$

where $H_d = 0$.

Equations 3 and 4 are extremely powerful, as they allow us to completely characterize the first two moments of the single-copy delay and number of hops, as described in Lemmas 1 and 2 below. By knowing the first two moments, we can use, for example, the moment matching approximation technique [23] to compute the approximate distribution of the expected delay.

Approximating the entire distribution of the delay rather than just computing, e.g, its expectation can be dramatically useful. For example, we can use it for establishing whether a message will reach its destination within a predefined time interval with a certain probability. This knowledge on the system can be more informative than just the simple average and can be used, e.g., to fine tune the system (for example, setting a TTL that allows messages to be delivered with a given probability).

In this work we focus our attention to the case in which nodes have exact knowledge. By exact knowledge we mean that nodes all know *exactly* the expected inter-meeting rate with their neighbours. Being all rate estimates $\hat{\mu}_{ij}$ exact (i.e., $\hat{\mu}_{ij} = \mu_{ij}$) during the forwarding process, all forwarding decisions are deterministic: a generic node i can identify with certainty who is a better next hop and thus to whom a message should be handed over. This implies that the forwarding probability $p_{ij}^{forw(\varphi)}$ (i.e., the probability that node i hands over the message to node j upon encounter) can be either 1 or 0. We refer the interested reader to [41] for the analysis of the imprecise knowledge case (which, at the expense of result accuracy, exploits approximation techniques not discussed here). In addition, please note that exact knowledge does not imply global knowledge: node i has only information on those nodes that it actually meets not on all the nodes of the network.

Under the exact knowledge assumption, we denote as $\mathcal{R}_i$ the set of suitable next hops for node i, defined as $\mathcal{R}_i = \{j : p_{ij}^{forw(\varphi)} > 0\}$. Probabilities $p_{ij}^{forw(\varphi)}$ can be computed directly from the definition of the forwarding strategy in use and are discussed for the reference forwarding policies in Section 5. Exploiting the definition of $\mathcal{R}_i$, in the lemmas below we describe how to compute the first and second moment of the delay and number of hops. Proofs for this section can be found in Appendix A.

Lemma 1 (Delay's first and second moment)**.** *The first and second moment of the delay D_i for a message generated by node i and addressed to node d can be obtained from the minimal non-negative solutions, if they exists, to the following systems, respectively:*

$$E[D_i] = E[T_i] + \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[D_j], \qquad \forall i \neq d \tag{5}$$

$$
\begin{aligned}
E[(D_i)^2] &= E[(T_i)^2] + \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[D_i^2] + \\
&+ 2 \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[T_{ij}] E[D_i], \qquad \forall i \neq d \qquad (6)
\end{aligned}
$$

where T_i is the time interval before the Markov chain exits from state i, T_{ij} is the time interval before the chain goes from state i to state j, and p_{ij} gives the probability of a transition from state i to state j.

Lemma 2 (Number of hops' first and second moment)**.** *The first and second moment of the number of hops H_i travelled by a message generated by node i and addressed to node d can be obtained, if they exists, from the minimal non-negative solutions to the following systems:*

$$
E[H_i] = 1 + \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_j], \qquad \forall i \neq d \qquad (7)
$$

$$
E[(H_i)^2] = \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_i^2] + 2 \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_i], \qquad \forall i \neq d \qquad (8)
$$

where p_{ij} denotes the probability of a transition from state i to state j in the Markov chain.

The two lemmas above can be solved once p_{ij}, T_i, and T_{ij} are known for all i, j pairs and for each of the forwarding policies in use. In the following we provide a general formulation for T_i, T_{ij}, and p_{ij}, which will be specialized later in the paper based on the distribution of inter-meeting times considered. Recalling that p_{ij} describes the probability that node i hands over the message to node j (equivalent to the probability that node j is the first possible forwarder encountered by node i), we can obtain probability p_{ij} as follows:

$$
p_{ij} = P\left(R_{ij} < \min_{z \in \mathcal{R}_i - \{j\}} \{R_{iz}\}\right), \quad j \in \mathcal{R}_i, \qquad (9)
$$

using the residual inter-meeting time R_{ij} to describe the time before the next encounter starting from message reception.

Let us now discuss how to compute T_{ij}, which describes the time before the chain in state i moves to state j (corresponding to a forwarding event between node i and node j). This event happens when node i meets j before meeting all other possible relays, i.e., when R_{ij} is smaller than all other

R_{iz}, with $z \in \mathcal{R}_i$. Thus, T_{ij} can be computed as the distribution of R_{ij} conditioned to be the minimum of $\{R_{iz}\}_{z\in\mathcal{R}_i}$. Thus, its PDF is given by:

$$P(T_{ij} = x) = \frac{P\left(R_{ij} = x, \{R_{iz} > x\}_{z\in\mathcal{R}_i-\{j\}}\right)}{\int_0^\infty P\left(R_{ij} = x, \{R_{iz} > x\}_{z\in\mathcal{R}_i-\{j\}}\right) \mathrm{d}x}, \tag{10}$$

obtained applying the law of conditional probability [42].

Finally, since T_i is defined as the time before node i hands over the message, T_i can be computed as the time before the first encounter with a possible forwarder, i.e., the following relation holds:

$$T_i = \min\{R_{ij}\}_{j\in\mathcal{R}_i}. \tag{11}$$

As evident from the above discussion, T_i, T_{ij}, and p_{ij} depend on i) the forwarding policy φ in use (see the dependence on $\mathcal{R}_i$) and ii) the distributions of inter-meeting times M_{ij}, which in turn characterize the distribution of residuals R_{ij}. Bullet i) is discussed in the next section, where the reference forwarding policies considered in this paper are introduced. Bullet ii) is discussed in Sections 6 and 8, where the general model presented above is specialized for the exponential, power law, and hyper-exponential inter-meeting times.

5. Reference Forwarding Strategies

Providing a model that is simple but at the same time complete enough to correctly describe the variety of existing single-copy forwarding approaches is not an easy task. In order to accomplish this goal, we abstract the variety of protocols described in Section 2 into the two main categories of social-oblivious (or blind) and social-aware forwarding protocols. For these categories, we consider the following policies, which identify important traits of existing forwarding strategies. More specifically, among the social-oblivious schemes we consider the following three policies.

Definition 3 (Direct Transmission)**.** *The source node can only deliver the message to the destination itself.*

Definition 4 (Always Forward)**.** *The source node hands over the message to the first node encountered, and so does each intermediate node. The process stops when the message is delivered to the destination.*

Definition 5 (Two Hop)**.** *The source node hands over the message to the first node encountered. If this first encounter is with the destination, the forwarding process is completed. Otherwise, the relay node is allowed to hand over the message only to the destination, if ever met.*

Such social-oblivious policies have been commonly used in the literature as baseline references [7] [27] [15]. The Direct Transmission and the Always Forward policies represent the two end points in the single-copy forwarding spectrum. The Two Hop scheme can be considered as an intermediate solution between these two extremes.

With regards to social-aware schemes, a common feature of all these algorithms is that a message (be it on the source node or on an intermediate relay) is handed over to another node only if the latter has a higher probability (we call it *fitness*) of bringing the message closer to its destination than the node currently holding the message. In the following, we consider fitness functions computed using only information on contacts between nodes, which have a direct dependence on the inter-meeting time distribution. This lets us clearly show what is the impact of the contact dynamics on the performance of opportunistic forwarding protocols. For the sake of completeness, in Appendix E we then discuss how the proposed analytical framework can be applied to more complex and popular social-aware policies, such as BUBBLE, SimBet, and HiBOp. Our two simplified reference social-aware policies are the following.

Definition 6 (Direct Acquaintance)**.** *The source and each intermediate relay hand over the message to the first encountered node having a higher fitness, where the fitness $f_{i,d}^{DA}$ of a generic node i for a message with destination d is defined as the estimated frequency $\hat{\mu}_{id}$ of a direct meeting with the destination d (Equation 12).*

$$f_{i,d}^{DA} = \hat{\mu}_{i,d}, \forall i \neq d \tag{12}$$

Definition 7 (Social Forwarding)**.** *Messages are delivered through a path with positive gradient of fitness, where the fitness $f_{i,d}^{SF}$ of node i for a message addressed to node d is computed (Equation 13) as the weighted sum of the fitness for a direct acquaintance ($f_{i,d}^{DA}$) and the fitness for an indirect meeting ($f_{i,d}^{I}$):*

$$f_{i,d}^{SF} = \beta f_{i,d}^{DA} + (1-\beta) f_{i,d}^{I}, \quad \textit{where } 0 < \beta < 1. \tag{13}$$

Component $f_{i,d}^{I}$ is a measure of the probability of being indirectly connected to the destination or, in other words, of the likelihood of being connected to *nodes that have high delivery probability for destination d. In the general case, it can be recursively defined as the weighted average of the social fitness of the encountered nodes, which implies:*

$$f^I_{i,d} = \sum_{j \in \mathcal{P}_i} w_{ij} \cdot \left(\gamma f^{DA}_{j,d} + (1-\gamma) f^I_{j,d}\right), \quad \textit{where } 0 < \gamma < 1. \tag{14}$$

In Equation 14, $\mathcal{P}_i$ denotes the set of nodes that can be encountered by node i, and $f^{DA}_{j,d}$ and $f^I_{j,d}$ are the direct and indirect fitness values of node i's neighbour j. Component $f^I_{j,d}$ ensures that high fitness values are also indirectly detected over multi-hop paths. We define w_{ij} as $\frac{\mu_{ij}}{\sum_{j \in \mathcal{P}_i} \mu_{ij}}$, thus w_{ij} weights the information about j based on the relative frequency of meeting j with respect to all other nodes. The rationale is that the information about j is as useful as node i is able to exploit it, i.e., as likely it is that node i can exploit node j as relay. Parameter γ is a weight that can be tuned in order to prioritize what neighbour j directly sees ($\gamma \to 1$, in this case) or what the neighbours of j see ($\gamma \to 0$, in this case). Parameter γ can be in general different from β in order to weight differently the fitness values associated directly with node i itself and those related to its neighbours. For the sake of simplicity, in the following we assume $\gamma = 1$.

Differently from the Direct Acquaintance policy, the Social Forwarding strategy is able to detect not only direct meetings with the destination, but also meetings with people that have a high probability of delivering the message to the destination. This strategy enables the exploitation of the delivery skills that are present in the environment surrounding the users, and not only of those of the user itself. In Section 7 we show how important it can be to exploit this feature.

If we assume a stationary mobility process and that nodes have an exact knowledge of the portion of the network they get in touch with (i.e., accurate information on their neighbourhood but no global knowledge, as discussed previously), nodes will be able to estimate with no error their expected inter-meeting rate with the other neighbours. Thus, when comparing its fitness value to that of another node, a generic node i will always make the same decision, either to forward, or not, to another node j. Thus, for each of the forwarding strategies defined above, we can compute the forwarding probability $p^{forw(\varphi)}_{ij}$ as specified in Table 2.

Table 2: Forwarding probability $p_{ij}^{forw(\varphi)}$

	DT	AF	2H	DA	SF
$p_{ij}^{forw(\varphi)}$	$\begin{cases} 1 & j=d \\ 0 & o.w. \end{cases}$	$\begin{cases} 0 & i=j \\ 1 & i \neq j \end{cases}$	$\begin{cases} 1 & i=s \vee (i \neq s \wedge j=d) \\ 0 & o.w. \end{cases}$	$\begin{cases} 1 & f_{i,d}^{DA} < f_{j,d}^{DA} \\ 0 & o.w. \end{cases}$	$\begin{cases} 1 & f_{i,d}^{SF} < f_{j,d}^{SF} \\ 0 & o.w. \end{cases}$

6. The model under different hypotheses for inter-meeting times

In Section 4 we have introduced a general framework for modeling the single-copy delays and number of hops, where this framework was dependent on T_i, p_{ij}, and T_{ij}. As discussed before, these quantities vary both with the specific forwarding strategy in use and with the specific distribution for inter-meeting times considered. While in Section 5 we discussed how to compute the forwarding probability $p_{ij}^{forw(\varphi)}$ for each of the chosen reference strategies, in the following we finally specialize T_i, p_{ij}, and T_{ij} to a set of relevant distributions for the inter-meeting times emerged from traces, exponential [18] and power law [15] inter-meeting times. Please note however that the proposed framework is more general than that, even if not in all cases closed form solutions are obtained. For example, when inter-meeting times feature a power law with exponential cut-off distribution [16], the model can be solved only numerically.

6.1. The Exponential Case

In this section we revisit the model proposed in Section 4 assuming that the inter-meeting time M_{ij} between a generic pair of nodes i, j is exponentially distributed with rate λ_{ij}. In this case, the rate μ_{ij} of the inter-meeting time exactly coincides with the rate λ_{ij} of the exponential distribution describing M_{ij}. Let us start our analysis with the computation of the time T_i required to exit state i for the chain in Figure 1.

Theorem 1 (Exit time)**.** *When inter-meeting times M_{ij} follow an exponential distribution with rate λ_{ij}, and forwarding scheme φ is in use, the time before the semi-Markov process exits state i follows an exponential distribution with rate $\sum_{j \in \mathcal{R}_i} \lambda_{ij}$. T_i's first two moments are thus given by the following:*

$$E[T_i] = \frac{1}{\sum_{j \in \mathcal{R}_i} \lambda_{ij}}, \tag{15}$$

$$E[T_i^2] = \frac{2}{[\sum_{j \in \mathcal{R}_i} \lambda_{ij}]^2}. \tag{16}$$

Proof. In order to apply Equation 11, which provide the formulation for T_i, we need to compute the distribution of R_{ij}, the residual inter-meeting time between node i and node j. Based on the memoryless property of the exponential distribution, we know that such R_{ij} follows an exponential distribution with the same rate λ_{ij}, i.e., $R_{ij} \sim Exp(\lambda_{ij})$. From standard probability theory we know that the minimum of a set of n exponential random variables is again a random variable with rate equal to the sum of the rates of the n random variables, obtaining $F_{T_i}(t) = e^{\sum_{j \in \mathcal{R}_i} \lambda_{ij} t}$. Then, the expectation and second moment of an exponential random variable follow directly (Equations 15 and 16). □

Theorem 1 proves that, under the exponential assumption for inter-meeting times, the semi-Markov process that describes the forwarding evolution becomes a Continuous Time Markov process, in which T_i follows an exponential distribution.

Below we derive the transition probabilities associated with the chain in Figure 1.

Theorem 2 (Transition probabilities p_{ij})**.** *When inter-meeting times M_{ij} follow an exponential distribution with rate λ_{ij}, and forwarding scheme φ is in use, transition probabilities p_{ij} for all $j \in \mathcal{R}_i$ are given by:*

$$p_{ij} = \frac{\lambda_{ij}}{\sum_{z \in \mathcal{R}_i} \lambda_{iz}}, \tag{17}$$

where λ_{ij} denotes the rate of encounters between node i and node j.

Proof. Recall (Equation 9) that $p_{ij} = P(R_{ij} < \min_{z \in \mathcal{R}_i - \{j\}}\{R_{iz}\})$. From standard probability theory we know that the minimum of a set of n exponential random variables is again a random variable with rate equal to the sum of the rates of the n random variables. Thus, $\min_z\{R_{iz}\} \sim Exp(\sum_z \lambda_{iz})$. Then, we have to compute the probability that R_{ij} is smaller than $\min_z\{R_{iz}\}$ ($P(R_{ij} < \min_z\{R_{iz}\}) = P(R_{ij} - \min_z\{R_{iz}\} < 0)$). This is a well known result from standard probability theory for exponential random variables and the solution is given in Equation 17. □

Finally, we consider T_{ij}. It is possible to prove the following theorem, which simply follows from substituting the expression for the PDF and CCDF of the residual inter-meeting times in Equation 10:

Theorem 3 (Exit time)**.** *When inter-meeting times M_{ij} follow an exponential distribution with rate λ_{ij}, and forwarding scheme φ is in use, T_{ij}, the time before the semi-Markov process goes from state i to state j, is distributed as T_i and its expected value is thus given by $\frac{1}{\sum_{j\in\mathcal{R}_i}\lambda_{ij}}$.*

Theorems 1, 2, and 3 completely define the forwarding Markov process in the case of inter-meeting times exponentially distributed. Thus, it is now straightforward to compute the expected delay and the expected number of hops travelled by messages using Lemmas 1 and 2.

6.2. The Power Law Case

In this section we revisit the analytical framework proposed in Section 4 when the inter-meeting times between a generic pair of nodes i and j follow a power law (Pareto[3]) distribution with shape α_{ij} and scale $t_{min_{ij}}$. In the following we use the definition of the Pareto distribution which allows for values arbitrarily close to zero and whose CCDF is shown in Equation 18. The expected value of such distribution is $\frac{t_{min_{ij}}}{\alpha_{ij}-1}$.

$$F_{M_{ij}}(t) = \left(\frac{t_{min_{ij}}}{t + t_{min_{ij}}}\right)^{\alpha_{ij}} \tag{18}$$

This version of the Pareto distribution is usually denoted as American Pareto [44]. We refer the interested reader to Appendix C for a throughout study of our analytical framework when the alternative definition of the Pareto distribution, usually denoted as European Pareto, is used. Please note that being the American Pareto a European Pareto shifted by $t_{min_{ij}}$ to the left, both Pareto definitions share the same requirements for having finite expectation, as discussed in more detail in [25]. Thus, the following remark holds.

Remark 1. *The Pareto distributions introduced above are defined for $\alpha_{ij} > 0$ (due to the required PDF normalization [45]), and their mean is finite when $\alpha_{ij} > 1$.*

Recall from Section 4 that T_i, T_{ij}, and p_{ij} are expressed in terms of the residual inter-meeting times R_{ij}, i.e., the time until the next contact between node i and node j starting from a random time t. Applying formula $F_{R_{ij}}(t) =$

[3]In the following we will use the terms "power law" and "Pareto" interchangeably.

$\frac{1}{E[M_{ij}]} \int_t^{\infty} F_{M_{ij}}(u)\mathrm{d}u$ [46], from an American Pareto random variable with shape α_{ij} and scale $t_{min_{ij}}$ we obtain residuals that feature an American Pareto distribution with shape $\alpha_{ij} - 1$ and scale $t_{min_{ij}}$. Similarly to the reference literature [15][16], for ease of computation in the following we restrict to the case of power law random variables having the same scale, i.e., $t_{min_{ij}} = t_{min}, \forall i, j$.

Remark 2. *The Pareto distribution of R_{ij} is defined for $\alpha_{ij} > 1$ (due to the required PDF normalization), and its mean is finite when $\alpha_{ij} > 2$.*

From a mathematical standpoint, Equations 9, 10, and 11 are mainly based on the computation of the minimum $\min_i X_i$ of a set of random variables $\{X_i\}_i$ and the computation of $P(X_1 < X_2)$, i.e., the probability that a random variable X_1 is smaller than another random variable X_2. When X_i features a Pareto distribution with shape α_i and scale t_{min} for all i values, it is possible to prove (see Appendix C) that $\min_i X_i$ follows a Pareto distribution with shape $\sum_i \alpha_i$ and scale t_{min}, while $P(X_1 < X_2)$ is equal to $\frac{\alpha_1}{\alpha_1+\alpha_2}$. Using these results, in Theorem 4 we derive the exit time T_i.

Theorem 4 (Exit time)**.** *When inter-meeting times M_{ij} follow a power law distribution with shape α_{ij} and scale t_{min} for all i, j pairs, and forwarding scheme φ is in use, the time T_i before the semi-Markov process exits state i follows a Pareto distribution with rate $\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n$ (where n denotes the cardinality $|\mathcal{R}_i|$ of the set $\mathcal{R}_i$) and scale t_{min}. From standard probability theory, the resulting first two moments of T_i, when finite, are thus:*

$$E[T_i] = \frac{t_{min}}{\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n - 1}, \tag{19}$$

$$E[T_i^2] = \frac{2(t_{min})^2}{(\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n)^2 - 3(\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n) + 2}. \tag{20}$$

Proof. The time before exiting state i is given by the time before handing over the message to any of the potential relays. This is equivalent to $T_i = \min_{j \in \mathcal{R}_i}\{R_{ij}\}$, which is power law distributed with shape $\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n$ and scale t_{min} (see Appendix C for the complete proof). Then Equation 20 follows from the definition of the first two moments of the Pareto distribution. □

Now we derive the transition probabilities of the Markov chain in Figure 1 under the power law assumption for inter-meeting times.

Theorem 5 (Transition probabilities p_{ij}). *When inter-meeting times M_{ij} are power law distributed with shape α_{ij} and scale t_{min}, and forwarding strategy φ is in use, transition probabilities p_{ij} are given by:*

$$p_{ij} = \frac{\alpha_{ij} - 1}{\sum_{z \in \mathcal{R}_i} \alpha_{iz} - n}, \tag{21}$$

where $n = |\mathcal{R}_i|$.

Proof. We know from Equation 9 that $p_{ij} = P(R_{ij} < \min\{R_{iz}\}_{z \in \mathcal{R}_i - \{j\}})$. As discussed in Appendix C, the minimum of power law random variables with the same scale is power law distributed with shape equal to the sum of the shapes of each random variable. Thus, p_{ij} reduces to the difference between two power law distributed random variables. Then, Equation 21 follows directly after applying the rule for deriving $P(X_1 < X_2)$ when both random variables feature a Pareto distribution, which we prove in Appendix C. □

Finally, we focus on T_{ij}. It is possible to prove the following theorem, which simply follows from substituting the expression for the PDF and CCDF of the residual inter-meeting times in Equation 10:

Theorem 6 (T_{ij}). *When inter-meeting times M_{ij} follow a power law distribution with shape α_{ij} and scale t_{min} for all i, j pairs, and forwarding scheme φ is in use, the time T_{ij} before the semi-Markov process goes from state i to state j is distributed as T_i, i.e., it follows a Pareto distribution with rate $\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n$ (where n denotes the cardinality $|\mathcal{R}_i|$ of the set $\mathcal{R}_i$) and scale t_{min}. Thus, its expectation is given by $E[T_i] = \frac{t_{min}}{\sum_{j \in \mathcal{R}_i} \alpha_{ij} - n - 1}$.*

The first and second moments of the delay and number of hops can now be computed from Lemmas 1 and 2.

7. Using the framework: two case studies

In this section we exemplify how the proposed framework can be used by discussing two case studies, and the performance of the Direct Transmission, Always Forward, Two Hop, Direct Acquaintance, and Social Forwarding schemes in such cases. These cases are exponential and power law inter-meeting times, which we have discussed in the previous section. Due to space limitations, here we focus only on the first moment of the delay and number of hops. Under the assumptions in Section 3 the proposed analytical model is exact, thus it is not compared with simulation results, which would simply generate totally overlapping curves.

In the following we consider 15 nodes, which move around in the network and exchange messages according to our reference forwarding policies. We consider the case of a heterogeneous network, in which we equally distribute our 15 nodes into 3 communities (hereafter denoted as $C1$, $C2$, and $C3$). We consider each community as being a complete subgraph, meaning that all nodes within each community share a social link with each other. We also add social links between nodes in different communities. As we assume that nodes' movements are triggered by their social relationships, these nodes will commute between different communities, and for this reason we denote them as *travellers*. This is an example of social-oriented mobility models, which are currently one of the most important approaches in the literature [47][48]. In the following, we consider two different scenarios, each of which is characterized by a different social structure connecting the nodes in different communities. More details on these social structures will be provided in the corresponding sections.

We define node mobility according to the following algorithm. For nodes that have only social relationships with members of their own community, we assume that each pair of nodes connected by a social link meets according to inter-meeting time M_{ij}, with default rate λ in the exponential case and default scale α in the power law case. If two nodes do not share a social link, they never get in touch with each other. For the sake of comparison, we want the expected inter-meeting time to be the same for both the exponential and power law case, thus we impose $\frac{1}{\lambda} = \frac{t_{min}}{\alpha-1}$. Without loss of generality, in the following we set t_{min} to 1 second, α to 3.5 (which guarantees finite expectation for both the inter-meeting times and their residuals) and, consequently, λ to $2.5s^{-1}$. For nodes that are connected with more than one community, we mimic the fact that the user divides its time between these groups by increasing its expected inter-meeting time with the members of these communities. So, basically, we keep constant the average number of peers encountered by each node, be it a traveller or a locally roaming user. Thus, for a generic node j that is in touch with n communities (or, equivalently, which is connected to nodes associated with n distinct communities), we force its expected inter-meeting times with any other node in those communities to be n times greater than that of another node i that is only connected with just one community. Thus, by imposing $\frac{1}{\lambda'} = n\frac{1}{\lambda}$ and

$\frac{t_{min}}{\alpha'-1} = n\frac{t_{min}}{\alpha-1}$, in the exponential case we have that the inter-meeting process of node j will be characterized by a rate λ' equal to $\frac{\lambda}{n}$, while in the power law case the scale α' for node j will be equal to $\frac{\alpha-1+n}{n}$.

For each of the reference forwarding schemes we plot the histogram of the expected delay and of the expected number of hops computed for any pair of nodes. In the case of 15 nodes, there are $n(n-1) = 210$ node pairs, for which we extract 210 values of expected delay and 210 values of expected number of hops solving the system of equations in Lemmas 1 and 2. The y-axis in all histograms shows the frequency of expected delay values normalized by the total number of expected delay samples (210, in this case). Bin width is chosen for each scenario in order to ensure the significance and readability of plots.

It is worth noting that results for the expected number of hops do not vary when the assumption about inter-meeting times is changed. This is due to our choice of keeping constant, within the same scenario, the expected inter-meeting times across the different distributions. In fact, as long as the expected inter-meeting times are the same, the relative ordering of delivery probabilities for each forwarding strategy, and thus the forwarding path, remains the same for both the exponential and power law case.

7.1. Scenario 1: travellers in each community

We start by considering the case of all three communities being directly connected by moving nodes. More specifically, focusing on community $C1$, we add one link connecting one node in $C1$ with one node in $C2$ and one link connecting one node in $C1$ with one node in $C3$. Using the same approach we connect one node in $C2$ to one node in $C1$ and one node in $C2$ to one node in $C3$, and the same is done for $C3$. As we assume that node movements are triggered by their social relationships with the other nodes of the network, community $C1$ will have two travellers visiting the other communities: specifically, one traveller goes to $C2$ and back, the other goes to $C3$ and back. The travellers in $C2$ and $C3$ have an analogous behaviour (Figure 2). This configuration ensures that the network is connected because it exists at least one multi-hop path between any pair of nodes. This allows us to show that, despite the network being connected, not all forwarding strategies are able to deliver messages between any node pair.

Figures 3 and 4 shows the forwarding performance as far as the delay is concerned. Specifically, we compute from the model the expected delay $E[D_{ij}]$ for all pairs i, j, and we plot in Figures 3 and 4 the distribution of

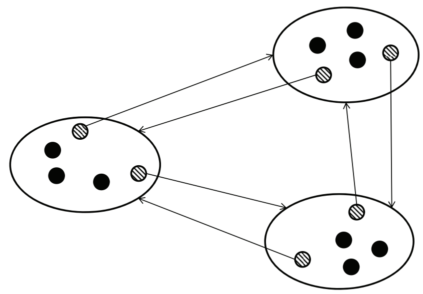

Figure 2: Scenario 1

the expected delay (across all pairs). The Direct Transmission scheme suffers when the source and the destination of the message do not get in touch with each other directly, thus producing infinite delays. This is because, with Direct Transmission, nodes can only deliver their messages directly to the destination, thus missing all the opportunities offered by relaying: when the destination is never met, the message cannot be delivered. However, relaying does not always guarantee a better performance in terms of expected delay, as the Two Hop case in Figures 3 and 4 shows. Recall that the expected delay is a weighted average of the expected delay of each possible path. Thus, if there exists even a single path with infinite expected delay, the overall expected delay will diverge. This is exactly what happens with the Two Hop strategy: due to the blind selection of the next hop, messages can take a wrong path at the first hop, and then they get stuck there because the intermediate relay node never meets the destination. In this scenario, such sequence of events is possible for all i, j source-destination pairs such that either (i) source node i and destination node j neither are traveller nor are in the same community or (ii) source node i is a traveller. In both cases there are some paths that achieve a finite expected delay, but there are also paths with infinite expected delay, and the latter drag the overall expected delay to infinite. Comparing the Two Hop scheme with the Direct Transmission strategy, in case (i) the fraction of node pairs that experience an infinite expected delay is the same under both protocols. In the second case, instead, i.e., when source node i is a traveller, among the possible paths that are added by the Two Hop scheme with respect to the Direct Transmission strategy, there are some characterized by an infinite delay, and those paths drag to infinite the expected delay for the Two Hop scheme, even if the direct encounter between the traveller and the destination would have a finite expectation. As an example of the first

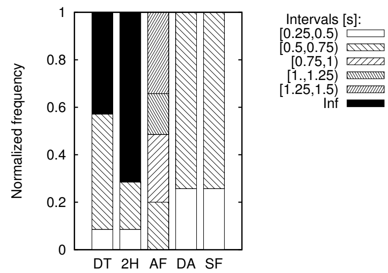

Figure 3: Distribution of the Expected Delay for Scenario 1 (exponential)

case, consider a message with source node in community $C1$ and destination node in community $C2$. In addition, assume that the source and destination nodes are not travellers. If the first encounter of the source node is with the traveller connecting $C1$ and $C3$, the message will be handed over to this node. However, this traveller never gets in touch directly with the destination in community $C2$, and the message will never be delivered. As for the second case, when the traveller is the source of the message (with destination in community $C1$, for example), there is always a non-negligible probability that, at the time the message is generated, the traveller is roaming in a community (C3, for example) different from the one in which the destination resides. In this case, the message will be handed over to the first encountered node, which, in our example, belongs to $C3$ and which will never meet the destination.

Direct Acquaintance, Social Forwarding, and Always Forward are able to exploit the social bridges between communities and to hand over the message to the convenient node. The Always Forward approach, however, forwards totally at random, and many hops may be required before the message eventually finds, by chance, its destination (Figure 5). Social strategies are instead able to choose the relays providing the best trade-off between low delay and efficient use of resources. Note also that in this scenario Direct Acquaintance and Social Forwarding show the same performance. In fact, they only differ when transitivity of contacts needs to be exploited for successful delivery, which is the case of the scenario discussed in the next section.

Please note that while the qualitative behavior is the same both in the case of exponential and power law inter-meeting times, in the latter case higher values of expected delays are more likely to appear due to the heavy tail of the power law distribution.

The expected delay and expected number of hops averaged across all node pairs are summarized in Tables 3 and 4.

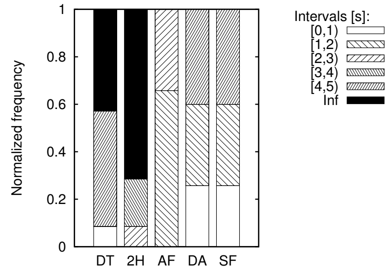

Figure 4: Distribution of the Expected Delay for Scenario 1 (power law)

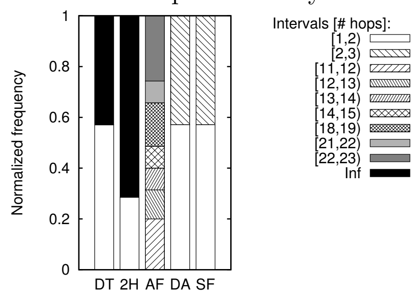

Figure 5: Distribution of the Expected Number of Hops for Scenario 1 (exp/power law)

7.2. Scenario 2: travellers in a single community

In this section we use the same scenario as in Section 7.1, except that we assign travellers only to community $C1$ (Figure 6). As in the previous case, the network is connected. However, while in Section 7.1 all communities were directly connected by means of traveller nodes, here $C2$ and $C3$ cannot communicate directly, and they have to exploit the forwarding capabilities of the visiting travellers from $C1$.

Figures 7 and 8 shows the expected delay experienced by messages in this scenario. The Direct Transmission, Two Hop, and Direct Acquaintance

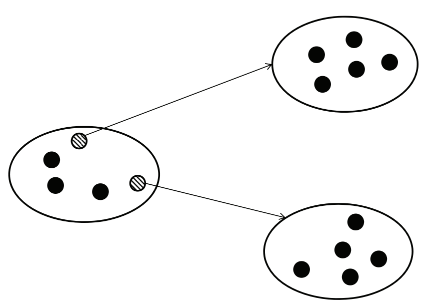

Figure 6: Scenario 2

	DT	2H	AF	DA	SF
exp	∞	∞	1.0481	0.546939	0.546939
pow	∞	∞	1.6112	2.28701	2.28701

Table 3: Average delay (s) – Scenario 1

	DT	2H	AF	DA	SF
exp/pow	∞	∞	16.8745	1.77143	1.77143

Table 4: Average number of hops – Scenario 1

schemes are not able to deliver a subset of messages. In the case of the Direct Transmission scheme the reason lies in the absence of direct contacts between the source of a message and its destination. The Two Hop scheme again suffers from the problem of messages that move away from their source node and get stuck at intermediate relays. In the case of the Direct Acquaintance policy, losses are due to the fact that a node hands over a message to another node that has a higher probability of meeting the destination, measured in terms of *direct* encounters only. The traveller that visits $C2$ does not meet any nodes of $C3$ directly, thus it is not considered a good relay for destinations in $C3$ by the Direct Acquaintance scheme. However, that traveller will meet in $C1$ the other traveller that visits $C3$ and thus it can be considered, indirectly, a good forwarder for $C3$ by nodes that roam only in $C2$. For this reason, a more efficient strategy should also consider the transitivity of opportunities (e.g., node a meets b, which in turn meets c, thus a can be considered a good relay for destination c). This transitivity of encounters is detected by the Social Forwarding strategy, which, for this reason, is able to deliver all messages to their destinations. The Always Forward strategy is, as before, able to deliver all messages, but using many relays (Figure 9), even more than in the previous scenario. The reason is that, being the forwarding opportunities so limited, with the Always Forward strategy the destination is typically found by chance after many (bad) relays have been used.

Note also that, as in the case analyzed in the previous section, while the qualitative behavior is the same both in the case of exponential and power law inter-meeting times, in the latter case higher values of expected delays are more likely to appear due to the power law heavy tail.

Summary results for the expected delay and the expected number of hops averaged across all node pairs are shown in Tables 5 and 6.

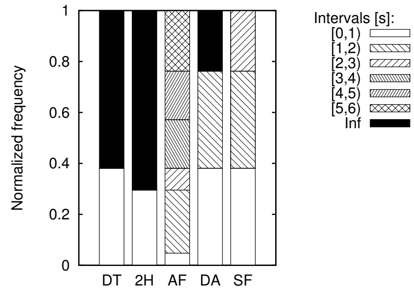

Figure 7: Distribution of the Expected Delay for Scenario 2 (exponential)

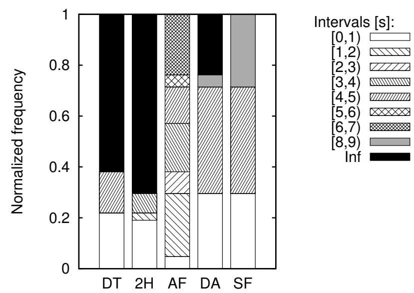

Figure 8: Distribution of the Expected Delay for Scenario 2 (power law)

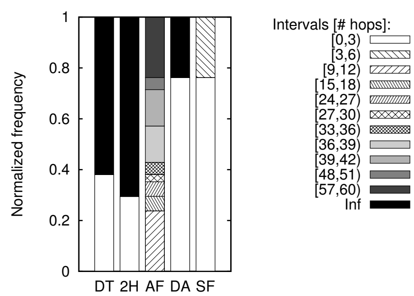

Figure 9: Distribution of the Expected Number of Hops for Scenario 2 (exp/power law)

8. Generalizing the framework

While in the section 6 we have solved exactly the model in the two very important cases of exponential and Pareto pairwise inter-meeting times, in this section we extend our analysis to a more general case. This case is represented by the class of positive random variables with coefficient of variation (the ratio between standard deviation and mean) greater than 1, usually referred to as high-variance distributions. These distributions are very relevant in the context of opportunistic networks, since they may imply a divergent

	DT	2H	AF	DA	SF
exp	∞	∞	3.35195	∞	1.2
pow	∞	∞	3.7167	∞	4.59114

Table 5: Average delay (s) – Scenario 2

	DT	2H	AF	DA	SF
exp/pow	∞	∞	35.1955	∞	2.35238

Table 6: Average number of hops – Scenario 2

expected delay. For this reason, in the following we focus on high-variance distributions.

A well-established technique (Lemma 3) for approximating high-variance distributions makes use of hyper-exponential distributions. A random variable X is hyper-exponentially distributed with n phases if X is, with probability p_i, an exponential random variable X_i with rate λ_i ($i = 1, \ldots, n$). Thus, the PDF of X is given by $\sum_{i=1}^{n} p_i f_{X_i}$, and the CCDF by $\sum_{i=1}^{n} p_i F_{X_i}$.

Lemma 3 (Hyper-exponential approximation)**.** *The two moments matching approximation of M_{ij} when the coefficient of variation ($c_{M_{ij}}$) is greater than 1 is a hyper-exponential random variable H_{ij} with $n = 2$ and parameters:*

$$p_1 = \frac{1}{2}\left(1 + \sqrt{\frac{c^2_{M_{ij}} - 1}{c^2_{M_{ij}} + 1}}\right) \quad , \quad p_2 = 1 - p_1, \tag{22}$$

$$\lambda_1 = \frac{2p_1}{E\left[M_{ij}\right]} \quad , \quad \lambda_2 = \frac{2p_2}{E\left[M_{ij}\right]}. \tag{23}$$

A throughout discussion on this technique can be found in [23]. Exploiting Lemma 3 we are then able to represent high-variance inter-meeting times with hyper-exponential random variables. Now we need to plug this hyper-exponential random variables into our framework, or, more specifically, to compute p_{ij}, T_i, and T_{ij} under the hyper-exponential case.

Let us start by stating some properties of hyper-exponential random variables, whose proofs can be found in Appendix D.

Lemma 4 (Residual)**.** *When the intermeeting time M_{ij} is hyper-exponentially distributed with n phases and parameters $p_1, \ldots, p_n$ and $\lambda_1, \ldots, \lambda_n$, the residual inter-meeting time R_{ij} follows again a hyper-exponential distribution with* *n phases, with CCDF:*

$$F_{R_{ij}}(t) = \sum_{i=1}^{n} r_i e^{-\lambda_i t}, \tag{24}$$

where $r_i = \frac{p_i \prod_{j \neq i} \lambda_j}{\sum_{z=1}^{n} p_z \prod_{j \neq z} \lambda_j}$.

Lemma 5 (Minimum)**.** *Consider m hyper-exponentially distributed random variables X_i (each with n_i phases and parameters $p_i^{(1)}, \ldots, p_i^{(n_i)}$ and $\lambda_i^{(1)}, \ldots, \lambda_i^{(n_i)}$). The minimum of these m random variables is hyper-exponentially distributed with CCDF:*

$$F_{min}(t) = \sum_{(z,\ldots,w) \in A_1 \times \ldots \times A_m} p_i^{(z)} \ldots p_j^{(w)} e^{-t(\lambda_i^{(z)} + \ldots + \lambda_j^{(w)})}, \tag{25}$$

where $A_i = \{1, \ldots, n_i\}$ and symbol $\times$ denotes the cartesian product.

Lemma 6 (Difference)**.** *Consider two hyper-exponentially distributed random variables X (phases n_X, probabilities l_i, rates λ_i) and Y (phases n_Y, probabilities m_i, rates μ_i). The probability that X is smaller than Y (equivalent to the CCDF of the difference $Y - X$) is given by the following:*

$$P(X < Y) = P(Y - X > 0) = 1 - \sum_{i=1}^{n_Y} m_i \mu_i \sum_{j=1}^{n_X} \frac{l_j}{\lambda_j + \mu_i}. \tag{26}$$

Exploiting the above lemmas, we can derive the following results.

Theorem 7 (Exit time)**.** *When inter-meeting times M_{ij} follow a hyper-exponential distribution with n_{ij} phases, probabilities $p_{ij}^{(k)}$ and rates $\lambda_{ij}^{(k)}$, and forwarding scheme φ is in use, the time T_i before the semi-Markov process exits state i follows a hyper-exponential distribution with $\prod_{j \in \mathcal{R}_i} n_{ij}$ phases, with probabilities and rates given by the following:*

$$p_u = r_{ij}^{(z)} * \ldots * r_{ik}^{(w)} \tag{27}$$

$$\lambda_u = \lambda_{ij}^{(z)} * \ldots * \lambda_{ik}^{(w)}, \tag{28}$$

for $j, k \in \mathcal{R}_i$, $(z, \ldots, w) \in \prod_{j \in \mathcal{R}_i} A_{ij}$, with $A_{ij} = \{1, \ldots, n_{ij}\}$, and $u \in \{1, \ldots, \prod_{j \in \mathcal{R}_i} n_{ij}\}$. From standard probability theory, the resulting first two moments of T_i are finite and equal to:

$$E[T_i] = \sum_{u=1}^{\prod_{j \in \mathcal{R}_i} n_{ij}} \frac{p_u}{\lambda_u} \tag{29}$$

$$E[T_i^2] = \sum_{u=1}^{\prod_{j \in \mathcal{R}_i} n_{ij}} \frac{2p_u}{\lambda_u^2}. \tag{30}$$

Proof. Recalling that $T_i = \min_{j \in \mathcal{R}_i}\{R_{ij}\}$, we use Lemma 4 for computing the residual inter-meeting times, then Lemma 5 for computing their minimum. □

Theorem 8 (Transition Probability)**.** *When inter-meeting times M_{ij} follow a hyper-exponential distribution with n_{ij} phases, probabilities $p_{ij}^{(k)}$ and rates $\lambda_{ij}^{(k)}$, and forwarding scheme φ is in use, transition probabilities p_{ij} are given by the following:*

$$p_{ij} = 1 - \sum_{w=1}^{\prod_{z \in \mathcal{R}_i - \{j\}} n_{iz}} m_w \mu_w \sum_{k=1}^{n_{ij}} \frac{p_{ij}^{(k)}}{\lambda_{ij}^{(k)} + \mu_w}, \tag{31}$$

where m_w and μ_w are the probabilities and rates of the hyper-exponential random variable describing the minimum $\min_{z \in \mathcal{R}_i - \{j\}}\{R_{iz}\}$ (which can be computed according to Lemma 5).

Proof. Recalling that $p_{ij} = P(R_{ij} < \min_{z \in \mathcal{R}_i - \{j\}}\{R_{iz}\})$, we use Lemma 4 for computing the residual inter-meeting times, then Lemma 5 for computing the minimum, and finally Lemma 6 for computing the difference □

Theorem 9 (T_{ij})**.** *When inter-meeting times M_{ij} follow a hyper-exponential distribution with n_{ij} phases, probabilities $p_{ij}^{(k)}$ and rates $\lambda_{ij}^{(k)}$, and forwarding scheme φ is in use, the time T_{ij} before the semi-Markov process goes from state i to state j is distributed as follows:*

$$f_{T_{ij}}(x) = \frac{\sum_{k=1}^{n_{ij}} p_{ij}^{(k)} e^{-\lambda_{ij}^{(k)} x} * \sum_{w=1}^{n_m} m_w e^{-\mu_w x}}{\sum_{w=1}^{n_m} m_w \mu_w \sum_{k=1}^{n_{ij}} \frac{p_{ij}^{(k)}}{\lambda_{ij}^{(k)} + \mu_w}}, \tag{32}$$

where random variable $\min_{z \in \mathcal{R}_i - \{j\}}$ is hyper-exponentially distributed with $n_m = \prod_{z \in \mathcal{R}_i - \{j\}} n_{iz}$ phases, probabilities m_w, and rates μ_w ($w = 1, \ldots, n_m$), which can be computed according to Lemma 5.

Proof. We apply Equation 10 using Lemma 4 for computing the residual inter-meeting times, then Lemma 5 for computing the minimum. □

Theorems 7, 8, and 9 completely define the forwarding Markov process in the case of inter-meeting times hyper-exponentially distributed. Thus, it is now straightforward to compute the expected delay and the expected number of hops travelled by messages using Lemmas 1 and 2. In addition, using the technique described in Lemma 3, we can use this approach to approximate the forwarding behavior under heterogeneous high-variance inter-meeting times.

9. Conclusion

In this paper we have proposed a general framework based on semi-Markov processes for modelling the forwarding process in opportunistic networks. Besides being independent of any specific forwarding policy, the framework is also independent of the specific hypothesis on the distribution of inter-meeting times between pairs of nodes, making it general enough to be used also when such hypothesis is changed. As an example, we have instantiated the framework in the two popular cases of exponential and Pareto inter-meeting times. In addition, we have used the model to compare the forwarding performance of social-oblivious and social-aware strategies in terms of expected delay and expected number of hops. Using this model, we have shown that social-aware policies in general provide lower delays while at the same time keeping the number of hops down, thus improving the efficiency of the network. Finally, we have discussed how, using the hyper-exponential distribution, the framework can be solved exactly and be used to approximate the forwarding behavior under high-variance inter-meeting times.

Acknowledgements

This work was partially funded by the European Commission under the SCAMPI (FP7-FIRE 258414), RECOGNITION (FET-AWARENESS 257756), EINS (FP7-FIRE 288021), and MOTO (FP7 317959) projects.

References

[1] S. Burleigh, A. Hooke, L. Torgerson, K. Fall, V. Cerf, B. Durst, K. Scott, H. Weiss, Delay-tolerant networking: an approach to interplanetary internet, IEEE Commun. Mag. 41 (6) (2003) 128–136.

[2] M. Conti, M. Kumar, Opportunities in opportunistic computing, IEEE Computer 43 (1) (2010) 42–50.

[3] J. Scott, P. Hui, J. Crowcroft, C. Diot, Haggle: A networking architecture designed around mobile users, in: WONS'06, 2006.

[4] A. Vahdat, D. Becker, Epidemic routing for partially connected ad hoc networks, Tech. rep. (2000).

[5] T. Spyropoulos, K. Psounis, C. Raghavendra, Efficient routing in intermittently connected mobile networks: The multiple-copy case, IEEE/ACM Trans. on Netw. 16 (1) (2008) 77–90.

[6] A. Lindgren, A. Doria, O. Schelén, Probabilistic routing in intermittently connected networks, LNCS (2004) 239–254.

[7] T. Spyropoulos, K. Psounis, C. Raghavendra, Efficient routing in intermittently connected mobile networks: The single copy case, IEEE/ACM Trans. on Netw. 16 (1) (2008) 63–76.

[8] C. Boldrini, M. Conti, A. Passarella, Exploiting users' social relations to forward data in opportunistic networks: The HiBOp solution, Pervasive and Mobile Computing 4 (5) (2008) 633–657.

[9] P. Hui, J. Crowcroft, E. Yoneki, Bubble rap: Social-based forwarding in delay tolerant networks, IEEE Trans. Mobile Comput. 10 (11) (2011) 1576 –1589.

[10] E. Daly, M. Haahr, Social network analysis for information flow in disconnected Delay-Tolerant MANETs, IEEE Trans. Mobile Comput. (2008) 606–621.

[11] X. Zhang, G. Neglia, J. Kurose, D. Towsley, Performance modeling of epidemic routing, Computer Networks 51 (10) (2007) 2867–2891.

[12] A. Al Hanbali, A. Kherani, P. Nain, Simple models for the performance evaluation of a class of two-hop relay protocols, in: IFIP NETWORKING'07, 2007, pp. 191–202.

[13] T. Spyropoulos, T. Turletti, K. Obraczka, Routing in Delay-Tolerant Networks Comprising Heterogeneous Node Populations, IEEE Trans. Mobile Comput. (2009) 1132–1147.

[14] C. Lee, D. Eun, Exploiting Heterogeneity in Mobile Opportunistic Networks: An Analytic Approach, in: IEEE SECON'10, IEEE, 2010, pp. 1–9.

[15] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, J. Scott, Impact of human mobility on opportunistic forwarding algorithms, IEEE Trans. Mobile Comput. (2007) 606–620.

[16] T. Karagiannis, J.-Y. Le Boudec, M. Vojnovic and, Power law and exponential decay of intercontact times between mobile devices, IEEE Trans. Mobile Comput. 9 (10) (2010) 1377 –1390. doi:10.1109/TMC.2010.99.

[17] V. Conan, J. Leguay, T. Friedman, Characterizing pairwise inter-contact patterns in delay tolerant networks, in: Autonomics'07, 2007.

[18] W. Gao, Q. Li, B. Zhao, G. Cao, Multicasting in delay tolerant networks: a social network perspective, in: ACM MobiHoc'09, ACM, 2009, pp. 299–308.

[19] X. Zhang, J. Kurose, B. N. Levine, D. Towsley, H. Zhang, Study of a Bus-Based Disruption Tolerant Network: Mobility Modeling and Impact on Routing, in: ACM Mobicom'07, 2007, pp. 195–206.

[20] V. Naumov, R. Baumann, T. Gross, An evaluation of inter-vehicle ad hoc networks based on realistic vehicular traces, in: ACM MobiHoc'6, 2006, pp. 108–119.

[21] A. Passarella, M. Conti, Characterising aggregate inter-contact times in heterogeneous opportunistic networks, in: NETWORKING 2011, 2011, pp. 301–313.

[22] H. Cai, D. Eun, Crossing over the bounded domain: From exponential to power-law intermeeting time in mobile ad hoc networks, IEEE/ACM Transactions on Networking (TON) 17 (5) (2009) 1578–1591.

[23] H. Tijms, J. Wiley, A first course in stochastic models, Vol. 2, Wiley Online Library, 2003.

[24] P. Tournoux, J. Leguay, F. Benbadis, V. Conan, M. De Amorim, J. Whitbeck, The accordion phenomenon: Analysis, characterization, and impact on dtn routing, in: INFOCOM 2009, IEEE, 2009, pp. 1116–1124.

[25] C. Boldrini, M. Conti, A. Passarella, Less is more: long paths do not help the convergence of social-oblivious forwarding in opportunistic networks, ACM/SIGMOBILE MobiOpp 12 (2012) 1–8.

[26] A. Al Hanbali, M. Ibrahim, V. Simon, E. Varga, I. Carreras, A survey of message diffusion protocols in mobile ad hoc networks, in: ValueTools '08, 2008, pp. 82:1–82:16.

[27] M. Grossglauser, D. Tse, Mobility increases the capacity of ad hoc wireless networks, IEEE/ACM Trans. on Netw. 10 (4) (2002) 477–486.

[28] Z. Haas, T. Small, A new networking model for biological applications of ad hoc sensor networks, IEEE/ACM Trans. on Netw. 14 (1) (2006) 27–40.

[29] H. Dubois-Ferriere, M. Grossglauser, M. Vetterli, Age matters: efficient route discovery in mobile ad hoc networks using encounter ages, in: ACM MobiHoc'03, ACM, 2003, pp. 257–266.

[30] V. Erramilli, A. Chaintreau, M. Crovella, C. Diot, Diversity of forwarding paths in pocket switched networks, in: ACM SIGCOMM'07, ACM, 2007, pp. 161–174.

[31] J. Silvis, D. Niemeier, R. D'Souza, Social networks and travel behavior: Report from an integrated travel diary, in: 11th International Conference on Travel Behaviour Research, Kyoto, 2006.

[32] A. Mtibaa, M. May, C. Diot, M. Ammar, Peoplerank: Social opportunistic forwarding, in: IEEE INFOCOM'10, 2010, pp. 1 –5. doi:10.1109/INFCOM.2010.5462261.

[33] P. Costa, C. Mascolo, M. Musolesi, G. Picco, Socially-aware routing for publish-subscribe in delay-tolerant mobile ad hoc networks, IEEE J. Sel. Areas Commun. 26 (5) (2008) 748 –760. doi:10.1109/JSAC.2008.080602.

[34] R. Groenevelt, P. Nain, G. Koole, The message delay in mobile ad hoc networks, Performance Evaluation 62 (1-4) (2005) 210–228.

[35] A. Al Hanbali, P. Nain, E. Altman, Performance of ad hoc networks with two-hop relay routing and limited packet lifetime, Performance Evaluation 65 (6-7) (2008) 463–483.

[36] A. Picu, T. Spyropoulos, T. Hossmann, An analysis of the information spreading delay in heterogeneous mobility dtns, in: World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2012 IEEE International Symposium on a, IEEE, 2012, pp. 1–10.

[37] A. Picu, T. Spyropoulos, Forecasting dtn performance under heterogeneous mobility: The case of limited replication, in: SECON 2012, IEEE, 2012, pp. 569–577.

[38] Y.-K. Ip, W.-C. Lau, O.-C. Yue, Performance modeling of epidemic routing with heterogeneous node types, in: IEEE ICC '08, 2008, pp. 219 –224. doi:10.1109/ICC.2008.48.

[39] U. Lee, S. Oh, K. Lee, M. Gerla, Scaling properties of delay tolerant networks with correlated motion patterns, in: ACM CHANTS'09, 2009, pp. 19–26.

[40] C. Boldrini, M. Conti, A. Passarella, Modelling social-aware forwarding in opportunistic networks, in: PERFORM'10, 2010, pp. 1–12.

[41] C. Boldrini, M. Conti, A. Passarella, Performance modelling of opportunistic forwarding with imprecise knowledge, in: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks (WiOpt), 2012 10th International Symposium on, IEEE, 2012, pp. 216–223.

[42] S. Ross, Introduction to probability models, Academic Press, 2007.

[43] M. Seshadri, S. Machiraju, A. Sridharan, J. Bolot, C. Faloutsos, J. Leskove, Mobile call graphs: beyond power-law and lognormal distributions, in: ACM SIGKDD'08, ACM, 2008, pp. 596–604.

[44] M. Rytgaard, Estimation in the Pareto distribution, ASTIN Bulletin 20 (2) (1990) 201 – 216.

[45] A. Clauset, C. Shalizi, M. Newman, Power-law distributions in empirical data, SIAM review 51 (4) (2009) 661–703.

[46] C. Boldrini, M. Conti, A. Passarella, From pareto inter-contact times to residuals, Communications Letters, IEEE (99) (2011) 1–3.

[47] C. Boldrini, A. Passarella, Hcmm: Modelling spatial and temporal properties of human mobility driven by users' social relationships, Computer Communications 33 (9) (2010) 1056–1074.

[48] M. Musolesi, C. Mascolo, Designing mobility models based on social network theory, ACM SIGMOBILE Mobile Computing and Communications Review 11 (3) (2007) 59–70.

Appendix A. Proofs for Section 4

Lemma 1 (Delay's first and second moment)**.** *The first and second moment of the delay D_i for a message generated by node i and addressed to node d can be obtained from the minimal non-negative solutions, if they exists, to the following systems, respectively:*

$$E[D_i] = E[T_i] + \sum_{j \neq d} p_{ij} E[D_j], \qquad \forall i \neq d$$

$$\begin{aligned} E[(D_i)^2] &= E[(T_i)^2] + \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[D_i^2] + \\ &+ 2 \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[T_{ij}] E[D_i], \qquad \forall i \neq d \end{aligned}$$

where T_i is the time interval before the Markov chain exits from state i, T_{ij} is the time interval before the chain goes from state i to state j, and p_{ij} gives the probability of a transition from state i to state j.

Proof. Let us start with the first moment. Exploiting the linearity of the expectation, we can rewrite Equation 3 as follows:

$$E[D_i] = \sum_{j \in \mathcal{R}_\rangle} p_{ij} E[T_{ij}] + \sum_{j \in \mathcal{R}_i - \{d\}} p_{ij} E[D_j] \tag{A.3}$$

Then, T_{ij} relates to the minimum T_i of the set of random variables $\{R_{ij}\}_{j \in \mathcal{R}_i}$ in the following way:

$$T_i = \left\{ \; T_{ij} \text{with probability } p_{ij}, \forall j \in \mathcal{R}_i \right. \tag{A.4}$$

Intuitively, T_i describes the time before source node i hands over the message to the first relay, while T_{ij} describes the distribution of R_{ij} knowing that i is the first relay. The n-th moment of T_i can thus be rewritten as follows:

$$E[(T_i)^n] = \sum_{j \in \mathcal{R}_i} p_{ij} E[(T_{ij})^n]. \tag{A.5}$$

In order to compute the second moment, we take the square of both sides of Equation 3:

$$(D_i)^2 = \begin{cases} (T_{id})^2 & p_{id} \\ (T_{ij} + D_j)^2 & p_{ij}, \forall j \in \mathcal{R}_i - \{d\} \end{cases}$$

After expansion, the above equation becomes:

$$\left(D_s^{(k)}\right)^2 = \begin{cases} \left(T_{id}\right)^2 & p_{id} \\ \left(T_{ij}\right)^2 + \left(D_j\right)^2 + 2T_{ij}D_j & p_{ij}, \forall j \in \mathcal{R}_i - \{d\} \end{cases}$$

Then, taking the expectation on both sides:

$$\begin{aligned} E[(D_i)^2] &= \sum_{j \in \mathcal{R}_i} p_{ij} E\left[(T_{ij})^2\right] + \\ &+ \sum_{j \in \mathcal{R}_i - \{d\}} p_{ij} E\left[D_j^2\right] + \\ &+ 2 \sum_{j \in \mathcal{R}_i - \{d\}} p_{ij} E\left[T_{ij}\right] E\left[D_j\right]. \end{aligned}$$

Equation 6 follows after substituting Equation A.5 in the above expression.
□

Lemma 2 (Number of hops' first and second moment)**.** *The first and second moment of the number of hops H_i travelled by a message generated by node i and addressed to node d can be obtained, if they exists, from the minimal non-negative solutions to the following systems:*

$$E[H_i] = 1 + \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_j], \qquad \forall i \neq d$$

$$E[(H_i)^2] = \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_i^2] + 2 \sum_{j \in \mathcal{R}_s - \{d\}} p_{ij} E[H_i], \qquad \forall i \neq d$$

where p_{ij} denotes the probability of a transition from state i to state j in the Markov chain.

Proof. The proof goes along the same line as the proof for Lemma 1 and we omit it. □

Appendix B. The European Pareto distribution

The notation presented in Section 6.2 is commonly referred to as American Pareto distribution. There exists also the European version of the power law distribution, which writes as follows:

$$F_E(t) = \left(\frac{t}{t_{min}}\right)^{-\alpha} \tag{B1}$$

Basically, being X a random variable following a European power law with scale t_{min} and scale α, then $Y = X - t_{min}$ is an American power law random variable.

Remark B1. *The expectation of a random variable featuring a European Pareto distribution with PDF defined as in Equation B1 is finite and equal to* $t_{min} \cdot \frac{\alpha}{\alpha-1}$ *when* $\alpha > 1$.

In order to apply the analytical model proposed in Section 4 to the case of inter-meeting times featuring a European Pareto distribution, we first need to compute the residual inter-meeting time, for which the following theorem holds (see [46] for the proof).

Theorem B1. *When inter-meeting time M features a European Pareto distribution with scale α and scale b ($F_M(t) = \left(\frac{b}{t}\right)^{\alpha}$), the residual inter-meeting time R is distributed as follows:*

$$F_R(t) = \begin{cases} \frac{t-\alpha t}{\alpha b} + 1 & t > 0 \wedge t \leq b \\ \frac{1}{\alpha}\left(\frac{b}{t}\right)^{-1+\alpha} & t > b \\ 0 & otherwise \end{cases} \tag{B2}$$

Remark B2. *The expectation of the residual of a European Pareto distribution with scale α is finite for all α values greater than* 2.

The analytical model proposed in Section 6.2 cannot be directly applied in this case. In fact, manipulating the residuals R_{ij} as we did for American power law inter-meeting times is not feasible, given that, for computing the minimum of the residuals, we would have to multiply the CCDFs in Equation B2 with each other. However, it is still possible to use an approximate model. In fact, it is straightforward to prove that $F_R(t) < \left(\frac{t}{t_{min}}\right)^{-(\alpha+1)}$, which is the PDF of a European Pareto random variable. By approximating the residual with a European Pareto random variable, we are able to use the analytical model discussed in Section 6.2.

Appendix C. Properties of power law distributions used in the paper

In this appendix we provide a general form for the minimum and difference between two power law distributed random variables. For the ease of computation, and without loss of generality, here we restrict to the case of power law random variables having the same scale, i.e., $t_{min_{ij}} = t_{min}, \forall i, j$. The following lemmas hold true both for the American and the European Pareto distribution.

Lemma C1 (Minimum of n Pareto Random Variables)**.** *The random variable X defined as $X = \min_i\{X_i\}$, where random variables X_i follow a power law distribution with scale α_i and scale x_{min}, is distributed according to a power law distribution with scale $\sum_i \alpha_i$.*

Proof. From standard probability theory we know that the CCDF of $\min_i\{X_i\}$ is equal to $\prod_i F_{X_i}$. When multiplying the CCDF of n power law random variables having the same scale, we again obtain a power law with scale equal to the sum of the scales of the n power law random variables. □

Remark C1. *The Pareto distribution resulting from the minimum of n Pareto distributions, each with its own scale α_i, is defined for $\sum_i \alpha_i > 0$ (due to the PDF normalization), and its mean is defined when $\sum_i \alpha_i > 1$.*

Lemma C2 (Comparison between two Pareto R.V.)**.** *Let us consider two random variables, X_1 and X_2, following a power law distribution with scale α_1 and α_2, respectively. Then, the probability that X_1 is lower than X_2 is given by:*

$$P(X_1 < X_2) = \frac{\alpha_1}{\alpha_1 + \alpha_2} \tag{C1}$$

Proof. We can rewrite $P(X_1 < X_2)$ using the law of total probability:

$$\begin{aligned} P(X_1 < X_2) &= \int_{x_{min}}^{+\infty} P(X_1 < X_2 | X_2 = y) P(X_2 = y) dy \\ &= \int_{x_{min}}^{+\infty} P(X_1 < y) P(X_2 = y) dy \end{aligned} \tag{C2}$$

Equation C1 is the solution to the above integral, computed after substituting the PDF and the CDF of the power law random variables into Equation C2. □

Appendix D. Properties of hyper-exponential distributions

Lemma 4 (Residual)**.** *When the intermeeting time M_{ij} is hyper-exponentially distributed with n phases and parameters $p_1, \dots, p_n$ and $\lambda_1, \dots, \lambda_n$, the residual inter-meeting time R_{ij} follows again a hyper-exponential distribution with* *n phases, with CCDF:*

$$F_{R_{ij}}(t) = \sum_{i=1}^{n} r_i e^{-\lambda_i t},$$

where $r_i = \frac{p_i \prod_{j \neq i} \lambda_j}{\sum_{z=1}^{n} p_z \prod_{j \neq z} \lambda_j}$.

Proof. The above result is obtained by simply solving the formula [46] for deriving the residual of a random variable, $F_{R_{ij}}(t) = \frac{1}{E[M_{ij}]} \int_t^{\infty} F_{M_{ij}}(u)\mathrm{d}u$. □

Lemma 5 (Minimum)**.** *Consider* m *hyper-exponentially distributed random variables* X_i *(each with* n_i *phases and parameters* $p_i^{(1)}, \ldots, p_i^{(n_i)}$ *and* $\lambda_i^{(1)}, \ldots, \lambda_i^{(n_i)}$*). The minimum of these* m *random variables is hyper-exponentially distributed with CCDF:*

$$F_{min}(t) = \sum_{(z,\ldots,w) \in A_1 \times \ldots \times A_m} p_i^{(z)} \ldots p_j^{(w)} e^{-t(\lambda_i^{(z)} + \ldots + \lambda_j^{(w)})},$$

where $A_i = \{1, \ldots, n_i\}$ *and symbol* $\times$ *denotes the cartesian product.*

Proof. It is a well-known result from probability theory that the CCDF of the minimum of m random variables is given by the product of the CCDF of each random variable. Thus, we can write the following:

$$F_{min}(t) = \prod_{i=1}^{m} \sum_{j=1}^{n_i} p_i^{(j)} e^{-t\lambda_i^{(j)}}. \tag{D21}$$

The above equation denotes the product of a set of m polynomials, which we can rewrite as in Equation 5. □

Lemma 6 (Difference)**.** *Consider two hyper-exponentially distributed random variables* X *(phases* n_X*, probabilities* l_i*, rates* λ_i*) and* Y *(phases* n_Y*, probabilities* m_i*, rates* μ_i*). The probability that* X *is smaller than* Y *(equivalent to the CCDF of the difference* $Y - X$*) is given by the following:*

$$P(X < Y) = P(Y - X > 0) = 1 - \sum_{i=1}^{n_Y} m_i \mu_i \sum_{j=1}^{n_X} \frac{l_j}{\lambda_j + \mu_i}.$$

Proof. Exploiting the law of total probability we can rewrite $P(X < Y)$ as follows:

$$\begin{aligned} P(X < Y) &= P(X < y|Y = y)P(Y = y) = \\ &= 1 - P(X > y|Y = y)P(Y = y) = \\ &= 1 - \int_0^\infty \left(\sum_{i=1}^{n_X} l_i e^{-\lambda_i y}\right) * \left(\sum_{j=1}^{n_Y} m_j \mu_j e^{-\mu_j y}\right) \mathrm{d}y \end{aligned}$$

Then, aften simple manipulation, Equation 6 follows. □

Appendix E. Modelling well known social-aware protocols

In Section 5 we touched on the ability of the proposed analytical framework to represent a variety of forwarding solutions. For the sake of completeness, here we discuss how the model can be applied to some well known social-aware policies proposed in the literature, specifically, BUBBLE [9], SimBet [10], and HiBOp [8]. Given the generality of the framework, it is sufficient to show how these algorithms can be mapped into appropriate definitions of the fitness of nodes as forwarders.

Appendix E.1. BUBBLE

The BUBBLE forwarding strategy is a combination of the LABEL and the RANK policies. In LABEL, nodes are assumed to be tagged with a label that identifies them as belonging to the same organization. A message is handed over upon encounter only if the peer shares the same label as the destination. According to this definition, the fitness of a node as a forwarder under the LABEL scheme is given by:

$$f_{i,d}^{LABEL} = \begin{cases} 1 & L(i) = L(d) \\ 0 & otherwise \end{cases}$$

where $L(i)$ gives node i's label. Under the RANK policy, messages are forwarded along a path of increasing node centrality. If we denote with c_i the node centrality of node i as defined in [9], we obtain the following:

$$f_{i,d}^{RANK} = c_i$$

In BUBBLE, the authors distinguish between global ranking and a local ranking, the latter being a node's centrality value with respect to the community it belongs to. Thus, we hereafter use $f_{i,d}^{RANK(global)}$ and $f_{i,d}^{RANK(local)}$

to differentiate the two rankings. The LABEL fitness and the RANK fitness (global and local) are then compared in order to select the best relay. More specifically, a message is forwarded to nodes with higher $f_{i,d}^{RANK(global)}$ as long as a no node belonging to the destination's community is found. Then, messages are handed over following an increasing path of $f_{i,d}^{RANK(local)}$.

Appendix E.2. SimBet

In SimBet [10], the fitness of a generic node i as a forwader for destination d is measured based on its ego-betweeness Bet_i and its similarity $Sim(i,d)$ with respect to the destination. The ego-betweeness expresses the centrality of the node in its ego network, while the similarity metric measures the number of common neighbors. We can now define f_i^{Bet} and $f_{i,d}^{Sim}$, as the fitness of node i according to its betweeness and its similarity to node d. f_i^{Bet} and $f_{i,d}^{Sim}$ can be computed directly from Equations 5 and 6 in [10]. We thus obtain:

$$f_{i,d}^{SimBet} = \alpha f_i^{Sim} + \beta f_{i,d}^{Bet}$$

Appendix E.3. HiBOp

The modelling of context-aware protocols like HiBOp [8] introduces additional complexity. So far, the fitness values have been depending only on node encounters, from which statistics on the meeting patterns or social network characteristics were extracted. On the contrary, in context-based forwarding protocols, nodes are enriched with a description of the environment the users operate in (e.g., the place they live, the company they work for, what they do in their leisure time) and this information is used to make more accurate predictions on the future encounters among nodes. Typically, the context is described by means of atomic pieces of information that we hereafter call *attributes*. Each attribute A_i takes a value from a set V_{A_i} of the possible values for that attribute. As an example, attribute `city` can take values `New York`, `Paris`, `Rome`, and so forth. The attribute values describing each node are collected in a table, called *Identity Table* (IT), which is exchanged upon contacts with other nodes. Using statistics on the neighbors' Identity Tables collected during pairwise meetings, nodes dynamically build their context-awareness and store this information into two other tables: the Current Context table contains information on the direct encounters, while the History table stores in an aggregate manner statistics on the context of the direct encounters. The overall forwarding fitness is then a composition of the fitness values computed for each of these tables, which we denote as $f_{i,d}^{IT}$,

$f_{i,d}^{CC}$, and $f_{i,d}^{H}$. Without providing further details on the way the protocols works (for which we refer the interested reader to [8]) we hereafter provide a convenient formulation for computing these fitness values. The Identity Table fitness is measured based on the correspondence between node i's IT and the destination's IT. Assuming that each IT is composed of K attributes, the IT fitness can be computed as follows:

$$f_{i,d}^{IT} = \frac{\sum_{k=1}^{K} w_k 1_{A_k(d)}(A_k(i))}{\sum_{k=1}^{K} w_k}.$$

$A_k(i)$ denotes the value of the k-th attribute in node i's IT and w_k the weight assigned to each attribute. The indicator function $1_{A_k(d)}(A_k(i))$ returns one when the value of the k-th attribute is the same in both node i's and node j's identity table, zero otherwise. The Current Context (CC) fitness can be computed as follows:

$$f_{i,d}^{CC} = \max_{j \in \mathcal{P}_i} f_{j,d}^{IT}$$

Finally, assuming $P_{op}^{(k)}$ gives the combination of the different statistics in the History tables as far as the k-th attribute is concerned, the History fitness can be computed as follows:

$$f_{i,d}^{H} = \frac{\sum_{k=1}^{K} w_k P_{op}^{(k)} 1_{A_k(d)}(A_k(i))}{\sum_{k=1}^{K} w_k},$$

where again $A_k(i)$ denotes the value of the k-th attribute in node i's IT, w_k the weight assigned to each attribute, and $1_{A_k(d)}(A_k(i))$ is an indicator function that returns one when the value of the k-th attribute is the same in both node i's and node j's identity table.